

Wed May '14 09:22:02 2003

us-09-914-831-1.rge

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 13, 2003, 15:11:51 ; Search time 1906 Seconds
(without alignments)
5725.893 Million cell updates/sec

Title: US-09-914-831-1
Sequence: 1 atgagcggtgagcagctgc.....cttgagcgaacagcgtac 375

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank :
1: gb_da :
2: gb_hg :
3: gb_in :
4: gb_lm :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vl :
15: em_da :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_ov :
22: em_or :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :
29: em_vl :
30: em_hg_hum :
31: em_hg_inv :
32: em_hg_other :
33: em_hg_mus :
34: em_hg_pin :
35: em_hg_rod :
36: em_hg_mam :
37: em_hg_vrt :
38: em_sy :
39: em_hgto_hum :
40: em_hgto_mus :
41: em_hgto_other :

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	375	100.0	375	5 AX035705	AX035705 Sequence
2	375	100.0	383	9 AF164795	AF164795 Homo sapi
3	375	100.0	604	9 BC024648	BC024648 Homo sapi
4	375	100.0	903	9 AF285119	AF285119 Homo sapi
5	375	100.0	1001	6 AX017845	AX017845 Sequence
6	370.2	98.7	636	9 HSM80069	AL136644 Homo sapi
7	359.8	95.9	1878	9 AY029161	AY029161 Homo sapi
8	286	76.3	1338	9 AF131857	AF131857 Homo sapi
9	270.8	72.2	581	10 BC028657	BC028657 Mus muscu
10	237.4	62.3	975	6 AX305959	AX305959 Sequence
11	160.8	42.9	182003	9 AL355987	AL355987 Human DNA
12	145	38.7	369	2 AC091284	AX410864 Sequence
13	109.8	29.3	123559	2 AC091284	AC091284 Mus muscu
14	109.8	29.3	236240	2 AL732590	AL732590 Mus muscu
15	98.2	26.2	101816	2 AC103205	AC103205 Rattus no
16	93.4	24.9	168384	2 AC121473	AC121473 Rattus no
17	89.4	23.8	330	6 AX410220	AX410220 Sequence
18	64	17.1	387	6 AX071589	AX071589 Sequence
19	59.2	15.8	704	3 AX071588	AX071588 Sequence
20	49.6	13.2	33428	8 AF332178	AF332178 Caenorhabd
21	46	12.3	650	8 AF332180	AF332180 Zea mays
22	43.4	11.6	1173	8 AF332180	AF332180 Zea mays
23	43	11.5	2967	3 GCW02	AF145597 Drosophill
24	42.2	11.3	625	3 AF145597	AF145597 Drosophill
25	42.2	11.3	184427	14 EH020824	U20824 Equine herp
26	42.2	11.0	5596	1 AF026544	AF026544 Rattus no
27	41.4	10.8	23069	1 AE008859	AE008859 Salimonei
28	40.6	10.7	1714	1 ABN1FHC	X51609 A.brazeiliens
29	40.2	10.6	1210	8 AF261275	AF261275 Oryza sat
30	39.8	10.6	1731	8 AF261275	AF261275 Oryza sat
31	39.8	10.6	1731	8 AF261275	AF261275 Oryza sat
32	39.8	10.6	127917	8 AC118673	AC118673 Genomic s
33	39.8	10.6	156933	8 AF485811	AC125411 Genomic s
34	39.8	10.6	162241	8 AF485811	AF485811 Oryza sat
35	39.4	10.5	33175	2 AC015165	AC015165 Drosophill
36	39.4	10.5	192350	3 AC008182	AC008182 Drosophill
37	39.4	10.5	197901	3 AC009523	AC009523 Drosophill
38	39.4	10.5	267547	3 AE003623	AE003623 Drosophill
39	39	10.4	714	3 AY013341	AY013341 Drosophill
40	39	10.4	720	3 AF393341	AF393341 Drosophill
41	39	10.4	39314	1 SGR300302	AF393341 Streptomy
42	39	10.4	152353	2 AC118777	AJ300302 Rattus no
43	39	10.4	176099	2 AC118777	AC118777 Rattus no
44	39	10.4	210727	2 AC122864	AC122864 Mus muscu
45	39	10.4	265050	1 AL627281	AL627281 Salimonei

ALIGNMENTS

RESULT 1	AX035705	375 bp	DNA	linear	PAT 15-NOV-2000
LOCUS	AX035705				
DEFINITION	Sequence 1 from Patent WO0052175.				
ACCESSION	AX035705				
VERSION	AX035705.1	GI:11191301			
KEYWORDS					
SOURCE					
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eultheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 375)				
AUTHORS	Kellner, R. and Klump, S.				
TITLE	Histidine protein-phosphatase				
JOURNAL	Patent: WO 0052175-A 1 08-SEP-2000;				

FEATURES MERCK PATENT GMBH (DE) ; KELLNER ROLAND (DE) ; KLUMPP SUSANNE (DE)
 source
 1. .375
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS
 1. .375
 /note="Human histidine protein phosphatase"
 /codon_start=1
 /protein_id="CAC16267.1"
 /db_xref="GI:1191302"
 /translation="MAVADLALIPVDIDSDGVFKYVILRVHSAPRSGAPAAESEIV
 RGYKAEVHADIVKVSQDMQKQDCDECIGSGRISHOSQDKIHVYGISNAYGPAQH
 AISTEKIKAKIPDEVTWANDY"
 BASE COUNT 83 a 110 c 112 g 70 t
 ORIGIN

Query Match
 Best Local Similarity 100.0%; Score 375; DB 6; Length 375;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGGCGACCTCGCTCATTCCTGATGTGAGCATCGACTCGACGGCGCTTC 60
 DB 1 ATGGCGGTGGGCGACCTCGCTCATTCCTGATGTGAGCATCGACTCGACGGCGCTTC 60
 QY 61 AAGTATGTGCTGATCCGAGTCCACTCGCTCCCGCTCGGGGGCTCCGGCTGCAAGAGC 60
 DB 61 AAGTATGTGCTGATCCGAGTCCACTCGCTCCCGCTCGGGGGCTCCGGCTGCAAGAGC 60
 QY 121 AAGGATGTGCTGCGCGGCTGACAAAGTGGGGTGTGATCCATGCGGACATCTACGAAAGTG 120
 DB 121 AAGGATGTGCTGCGCGGCTGACAAAGTGGGGTGTGATCCATGCGGACATCTACGAAAGTG 120
 QY 181 TCGGGGACATGACAGAAAGAGGCTGGGCTGAGTGTGTGGGGCGGGCGCATCTCC 240
 DB 181 TCGGGGACATGACAGAAAGAGGCTGGGCTGAGTGTGTGGGGCGGGCGCATCTCC 240
 QY 241 CACGAGATCGAGGAGAAAGATTCACGTGTAGCGGCTATTCATGCGCTATGCTCTGCG 300
 DB 241 CACGAGATCGAGGAGAAAGATTCACGTGTAGCGGCTATTCATGCGCTATGCTCTGCG 300
 QY 301 CAGCAGCCGATTTCACTGAGAAATCAAGCAATATACCCGAGCTAGAGGTCACTGG 360
 DB 301 CAGCAGCCGATTTCACTGAGAAATCAAGCAATATACCCGAGCTAGAGGTCACTGG 360
 QY 361 GCTAACGACGGCTAC 375
 DB 361 GCTAACGACGGCTAC 375

RESULT 2
 AF164795 583 bp mRNA linear pri 02-JUL-2000
 LOCUS AF164795
 DEFINITION Homo sapiens sex-regulated protein janus-a mRNA, complete cds.
 ACCESSION AF164795
 VERSION AF164795.1 GI:8895092
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 583)
 Li, Y., Shi, J., Huang, C., Jiang, C., Ren, S., Zhou, J., Yu, Y., Xu, S.,
 Wang, Y., Fu, G., Chen, Z. and Han, Z.
 A novel gene expressed in human adrenal gland
 Unpublished
 2 (bases 1 to 583)
 Peng, Y., Gu, Y., Li, Y., Fu, S., Gu, J., Gu, W., Jiang, C., Yu, Y.,
 Han, Z., Wang, Y., Chen, Z. and Fu, G.
 Direct Submission
 Submitted (02-JUL-1999) Chinese National Human Genome Center at
 Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
 Shanghai 201203, P. R. China
 Location/Qualifiers

source
 1. .583
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="adrenal gland"
 CDS
 9. .386
 /note="JANA"
 /codon_start=1
 /evidence="not_experimental"
 /product="sex-regulated protein janus-a"
 /protein_id="AA80759.1"
 /db_xref="GI:8895093"
 /translation="MAVADLALIPVDIDSDGVFKYVILRVHSAPRSGAPAAESEIV
 RGYKAEVHADIVKVSQDMQKQDCDECIGSGRISHOSQDKIHVYGISNAYGPAQH
 AISTEKIKAKIPDEVTWANDY"
 BASE COUNT 133 a 181 c 164 g 105 t
 ORIGIN

Query Match
 Best Local Similarity 100.0%; Score 375; DB 9; Length 583;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGGCGACCTCGCTCATTCCTGATGTGAGCATCGACTCGACGGCGCTTC 60
 DB 9 ATGGCGGTGGGCGACCTCGCTCATTCCTGATGTGAGCATCGACTCGACGGCGCTTC 60
 QY 61 AAGTATGTGCTGATCCGAGTCCACTCGCTCCCGCTCGGGGGCTCCGGCTGCAAGAGC 68
 DB 61 AAGTATGTGCTGATCCGAGTCCACTCGCTCCCGCTCGGGGGCTCCGGCTGCAAGAGC 68
 QY 121 AAGGATGTGCTGCGCGGCTGACAAAGTGGGGTGTGATCCATGCGGACATCTACGAAAGTG 120
 DB 121 AAGGATGTGCTGCGCGGCTGACAAAGTGGGGTGTGATCCATGCGGACATCTACGAAAGTG 120
 QY 181 TCGGGGACATGACAGAAAGAGGCTGGGCTGAGTGTGTGGGGCGGGCGCATCTCC 240
 DB 181 TCGGGGACATGACAGAAAGAGGCTGGGCTGAGTGTGTGGGGCGGGCGCATCTCC 240
 QY 241 CACGAGATCGAGGAGAAAGATTCACGTGTAGCGGCTATTCATGCGCTATGCTCTGCG 300
 DB 241 CACGAGATCGAGGAGAAAGATTCACGTGTAGCGGCTATTCATGCGCTATGCTCTGCG 300
 QY 301 CAGCAGCCGATTTCACTGAGAAATCAAGCAATATACCCGAGCTAGAGGTCACTGG 360
 DB 301 CAGCAGCCGATTTCACTGAGAAATCAAGCAATATACCCGAGCTAGAGGTCACTGG 360
 QY 361 GCTAACGACGGCTAC 375
 DB 361 GCTAACGACGGCTAC 375

RESULT 3
 BC024648 604 bp mRNA linear pri 12-MAR-2002
 LOCUS BC024648
 DEFINITION Homo sapiens, similar to HSPC141 protein, clone MGC:24855
 ACCESSION BC024648
 VERSION BC024648.1 GI:19353099
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 604)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAR-2002) National Institutes of Health, Mammalian
 Genome Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.

REMARK COMMENT

CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amgdbcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huliyil, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 29 Row: 0 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7661799.
 Location/Qualifiers

FEATURES

SOURCE

1..604
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:24855 IMAGE:4934238"
 /tissue_type="Brain, anaplastic oligodendroglioma with
 1p/19q loss"
 /clone_id="NCL CGAP_Brn67"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORE6"
 22..399
 /codon_start=1
 /product="Similar to HSPC141 protein"
 /protein_id="AAH24648.1"
 /db_xref="gi:19353100"
 /translation="MAVADLALIPVDIDSDGVKRYVLIIRHSPAGAPAESEKIV
 RGYKWEYHADIYKVSQDMOKGDCBELGGGRISHOSODKKIHVYGSMAVGPAGH
 AISTEKIKAKTPEYEWANDGY"
 AISTEKIKAKTPEYEWANDGY"

CDS

BASE COUNT 145 a 182 c 171 g 106 t
 ORIGIN

Query Match 100.0%; Score 375; DB 9; Length 604;
 Best Local Similarity 100.0%; Pred. No. 2.6e-75;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGTGACATCGACTCCGAGCGCTTC 60
 |||||||
 DB 22 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGTGACATCGACTCCGAGCGCTTC 81
 |||||||
 OY 61 AAGTATGTCTGATCGAGTCCAGTCCGCTCCGCGGCTCCGCGTCCGAGAGC 120
 |||||||
 DB 82 AAGTATGTCTGATCGAGTCCAGTCCGCTCCGCGGCTCCGCGTCCGAGAGC 141
 |||||||
 OY 121 AAGGAGATGTCGCGGCTACAGTGGGCTGAGTACCATCGGACATCTACGAAAGTG 180
 |||||||
 DB 142 AAGGAGATGTCGCGGCTACAGTGGGCTGAGTACCATCGGACATCTACGAAAGTG 201
 |||||||
 OY 181 TCGGGGACATGCAAGAGCAAGCTGCACTGTGAGTGTCTGGGCGGCGGCGCATCTCC 261
 |||||||
 DB 202 TCGGGGACATGCAAGAGCAAGCTGCACTGTGAGTGTCTGGGCGGCGGCGCATCTCC 300
 |||||||
 OY 241 CACCAAGTCAAGCAAGAGCAAGTACGCTATTCATGCTATGCTCTGCC 321
 |||||||
 DB 262 CACCAAGTCAAGCAAGAGCAAGTACGCTATTCATGCTATGCTCTGCC 342
 |||||||
 OY 301 CAGCAGCGCATTTCACTGAGAAATCAAAAGCAAGTACCGGATAGAGTCACTGG 360
 |||||||
 DB 322 CAGCAGCGCATTTCACTGAGAAATCAAAAGCAAGTACCGGATAGAGTCACTGG 381
 |||||||
 OY 361 GCTAACGACGGCTAC 375
 |||||||
 DB 382 GCTAACGACGGCTAC 396
 |||||||

RESULT 4
 AF285119 AF285119 903 bp mRNA linear PRI 21-FEB-2001
 LOCUS

DEFINITION Homo sapiens CGI-202 mRNA, complete cds.
 ACCESSION AF285119
 VERSION AF285119.1 GI:9858828
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Lai, C.H., Chiu, J.Y., and Lin, W.
 TITLE Identification of the human crooked neck gene by comparative gene
 identification
 JOURNAL Biochim. Biophys. Acta 1517 (3), 449-454 (2001)
 MEDLINE 21240337
 PUBMED 11342225

REFERENCE
 AUTHORS Lin, W.-C.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2000) Institute of Biomedical Sciences, Academia
 Sinica, No. 128, Section 2, Academia Road, Taipei, Taiwan 115,
 Taiwan

FEATURES

SOURCE

1..903
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 334..711
 /codon_start=1
 /product="CGI-202"
 /protein_id="AA01156.1"
 /db_xref="gi:9858829"
 /translation="MAVADLALIPVDIDSDGVKRYVLIIRHSPAGAPAESEKIV
 RGYKWEYHADIYKVSQDMOKGDCBELGGGRISHOSODKKIHVYGSMAVGPAGH
 AISTEKIKAKTPEYEWANDGY"
 AISTEKIKAKTPEYEWANDGY"

CDS

BASE COUNT 170 a 290 c 270 g 173 t
 ORIGIN

Query Match 100.0%; Score 375; DB 9; Length 903;
 Best Local Similarity 100.0%; Pred. No. 2.5e-75;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGTGACATCGACTCCGAGCGCTTC 60
 |||||||
 DB 334 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGTGACATCGACTCCGAGCGCTTC 93
 |||||||
 OY 61 AAGTATGTCTGATCGAGTCCAGTCCGCTCCGCGGCTCCGCGTCCGAGAGC 453
 |||||||
 DB 394 AAGTATGTCTGATCGAGTCCAGTCCGCTCCGCGGCTCCGCGTCCGAGAGC 513
 |||||||
 OY 121 AAGGAGATGTCGCGGCTACAGTGGGCTGAGTACCATCGGACATCTACGAAAGTG 180
 |||||||
 DB 454 AAGGAGATGTCGCGGCTACAGTGGGCTGAGTACCATCGGACATCTACGAAAGTG 240
 |||||||
 OY 181 TCGGGGACATGCAAGAGCAAGCTGCACTGTGAGTGTCTGGGCGGCGGCGCATCTCC 261
 |||||||
 DB 514 TCGGGGACATGCAAGAGCAAGCTGCACTGTGAGTGTCTGGGCGGCGGCGCATCTCC 300
 |||||||
 OY 241 CACCAAGTCAAGCAAGAGCAAGTACGCTATTCATGCTATGCTCTGCC 321
 |||||||
 DB 574 CACCAAGTCAAGCAAGAGCAAGTACGCTATTCATGCTATGCTCTGCC 342
 |||||||
 OY 301 CAGCAGCGCATTTCACTGAGAAATCAAAAGCAAGTACCGGATAGAGTCACTGG 360
 |||||||
 DB 634 CAGCAGCGCATTTCACTGAGAAATCAAAAGCAAGTACCGGATAGAGTCACTGG 693
 |||||||
 OY 361 GCTAACGACGGCTAC 375
 |||||||
 DB 694 GCTAACGACGGCTAC 708
 |||||||

RESULT 5
 AX017845 AX017845 1001 bp DNA linear PAT 07-SEP-2000
 LOCUS

DEFINITION Sequence 21 from Patent WO9946375.
 AX017845
 VERSION AX017845.1 GI:10042448
 KEYWORDS
 SOURCE
 ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and
 Pilsarsky, C.
 Human nucleic acid sequences from prostate tissue
 Patent: WO 9946375-A 21.16-SEP-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERNARD (DE); ROSENTHAL ANDRE (DE); METAGIN GES FUER GENOMFORSCHUN
 (DE); PILSARSKY CHRISTIAN (DE)

FEATURES
 Location/Qualifiers
 1..1001

BASE COUNT 178 a 312 c 329 g 182 t
 ORIGIN

Query Match
 Best Local Similarity 100.0%; Score 375; DB 6; Length 1001;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGCGGACCTCGCTCTCAATCCGTATGTGGACATCGACTCCAGCGGCTCTTC
 Db 435 ATGGCGGTGGCGGACCTCGCTCTCAATCCGTATGTGGACATCGACTCCAGCGGCTCTTC
 QY 61 AAGTATGTGCTGATCCGAGTCCATCGGCTCCCGCTCCGCGGCTCCGCGCTCAGAGAGC
 Db 495 AAGTATGTGCTGATCCGAGTCCATCGGCTCCCGCTCCGCGGCTCCGCGCTCAGAGAGC
 QY 121 AAGGATGTCGTCGGCGCTACCAAGTGGCTGATACATCGGACATCTAGACAAAGTG
 Db 555 AAGGATGTCGTCGGCGCTACCAAGTGGCTGATACATCGGACATCTAGACAAAGTG
 QY 181 TCGGGCGCATGCAAGCAAGAGCTCGACTGATGTCGTCGGCGGCGGCGCATCTCC
 Db 615 TCGGGCGCATGCAAGCAAGAGCTCGACTGATGTCGTCGGCGGCGGCGCATCTCC
 QY 241 CACGAGTGCAGGACAAAGATTCACGTGAGGCTATTCATGAGGCTATGCTCTGCC
 Db 675 CACGAGTGCAGGACAAAGATTCACGTGAGGCTATTCATGAGGCTATGCTCTGCC
 QY 301 CACGAGCCATTTCAACTGAGAAATCAAGCAAGCAAGTACCCGACTAGAGAGTCACTGG
 Db 735 CACGAGCCATTTCAACTGAGAAATCAAGCAAGTACCCGACTAGAGAGTCACTGG
 QY 361 GCTAACGAGGCTAC 375
 Db 795 GCTAACGAGGCTAC 809

RESULT 6
 HSW80069
 LOCUS HSW80069 636 bp mRNA linear PRI 12-JUL-2002
 DEFINITION Homo sapiens mRNA: cDNA DKFZp564M173 (from clone DKFZp564M173);
 complete cds.
 ACCESSION AL136644
 VERSION AL136644.1 GI:12052813
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 Direct Submission
 Submitted (10-JUL-2002) GSF, Institut fuer Bioinformatik MIPS,
 Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by ACOMA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFZp564M173) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

Location/Qualifiers
 1..636

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="DKFZp564M173"
 /tissue_type="brain"
 /clone_lib="564 (synonym: hfbz2). Vector pAMP1, host
 XI-2Blue; sites NotI + SalI"
 /dev_stage="fetal"
 1..636
 /gene="DKFZp564M173"
 26..403
 /gene="DKFZp564M173"
 /note="similarity to Drosophila janus A"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA66579.1"
 /db_xref="GI:12052814"
 /db_xref="SPTRMBL:O9H0Y3"
 /translation="NAVDALIPDNDSDGVFKYILINISAPSGAPAESEIY
 RGYMAEYHADIYDRVSGMDKQKDCDEJAGGTSQSDKRIHYGYSMAYRAQH
 AISTEKIKKPYDEYTMANGY"
 561..566
 /gene="DKFZp564M173"
 575

BASE COUNT 176 a 184 c 172 g 104 t
 ORIGIN

Query Match
 Best Local Similarity 98.7%; Score 370.2; DB 9; Length 636;
 Matches 372; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGCGGACCTCGCTCTCAATCCGTATGTGGACATCGACTCCAGCGGCTCTTC
 Db 26 ATGGCGGTGGCGGACCTCGCTCTCAATCCGTATGTGGACATCGACTCCAGCGGCTCTTC
 QY 61 AAGTATGTGCTGATCCGAGTCCATCGGCTCCCGCTCCGCGGCTCCGCGCTCAGAGAGC
 Db 86 AAGTATGTGCTGATCCGAGTCCATCGGCTCCCGCTCCGCGGCTCCGCGCTCAGAGAGC
 QY 121 AAGGATGTCGTCGGCGCTACCAAGTGGCTGATACATCGGACATCTAGACAAAGTG
 Db 146 AAGGATGTCGTCGGCGCTACCAAGTGGCTGATACATCGGACATCTAGACAAAGTG
 QY 181 TCGGGCGCATGCAAGCAAGAGCTCGACTGATGTCGTCGGCGGCGGCGCATCTCC
 Db 206 TCGGGCGCATGCAAGCAAGAGCTCGACTGATGTCGTCGGCGGCGGCGCATCTCC
 QY 241 CACGAGTGCAGGACAAAGATTCACGTGAGGCTATTCATGAGGCTATGCTCTGCC
 Db 266 CACGAGTGCAGGACAAAGATTCACGTGAGGCTATTCATGAGGCTATGCTCTGCC
 QY 301 CACGAGCCATTTCAACTGAGAAATCAAGCAAGCAAGTACCCGACTAGAGAGTCACTGG
 Db 326 CACGAGCCATTTCAACTGAGAAATCAAGCAAGTACCCGACTAGAGAGTCACTGG
 QY 361 GCTAACGAGGCTAC 375
 Db 386 GCTAACGAGGCTAC 400

RESULT 7

Yy	373	TAC 375
Db	1676	TAC 1678
RESULT 8		
AF131857		1338 bp mRNA linear PRI 12-MAR-1999
LOCUS		
DEFINITION	Homo sapiens clone 24870 mRNA sequence.	
ACCESSION	AF131857	
VERSION	AF131857.1 GI:4406704	
KEYWORDS	PII CDNA.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1338) Andersson, B., Wertheim, M.A., Ricafrente, J.Y., Liu, W. and Gibbs, R.A. A 'double adaptor' method for improved shotgun library construction Anal. Biochem. 236 (1), 107-113 (1996)	
AUTHORS	2 (bases 1 to 1338) Yu, W., Andersson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W., Ricafrente, J.Y., Wertheim, M.A., Lennon, G. and Gibbs, R.A. Large-scale concatenation cDNA sequencing Genome Res. 7 (4), 353-358 (1997)	
JOURNAL	97264341	
MEDLINE	9710174	
PUBMED	3 (bases 1 to 1338) Mei, G., Yu, W. and Gibbs, R.A. Direct Submission Submitted (26-FEB-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza RM N152L, Houston, TX 77030, USA Contact gme@bcm.tmc.edu for more information. Location/Qualifiers 1..1338 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="24870" /note="The I.M.A.G.E. Consortium clone ID member is 24870 and the library (INTB) was derived from human female infant brain tissue."	
REMARK		
FEATURES		
ORIGIN		
BASE COUNT	249 a 455 c 396 g 238 t	
Query Match	76.3%; Score 286; DB 9; Length 1338;	
Best Local Similarity	100.0%; Pred. NO. 4; e-55; Indels 0; Gaps 0;	
Matches 286; Conservative		
OY	1 ATGGCGGTGCGAGCCGCTCATTCCTGATGTGACATGCATCCTCGAGCGGGTCTTC 60	
Db	340 ATGGCGGTGCGAGCCGCTCATTCCTGATGTGACATGCATCCTCGAGCGGGTCTTC 399	
OY	61 AAGTAGTGCTGATTCGACGTCACATCGGGCTCCCGCTCCGGGGCTCCGGCTCGACAGAGC 120	
Db	400 AAGTAGTGCTGATTCGACGTCACATCGGGCTCCCGCTCCGGGGCTCCGGCTCGACAGAGC 459	
OY	121 AAGGAGATCGTCGGCGGCTACCAATGGGCTGAATACCATTGGCGGACATCGACACAAGTG 180	
Db	460 AAGGAGATCGTCGGCGGCTACCAATGGGCTGAATACCATTGGCGGACATCGACACAAGTG 519	
OY	181 TCGGGCGACATCGACAGAACAGGCTGGAGACTGTGTGGCGCGGGCGGACATCTCC 240	
Db	520 TCGGGCGACATCGACAGAACAGGCTGGAGACTGTGTGGCGCGGGCGGACATCTCC 579	
OY	241 CACCAAGTTCAGGACAAAGAATTTCACTGTACGGCTATTCCATTCC 286	
Db	580 CACCAAGTTCAGGACAAAGAATTTCACTGTACGGCTATTCCATTCC 625	
RESULT 9		

BC028657
LOCUS 581 bp
DEFINITION Mus musculus, RIKEN cDNA 170008c22 gene, clone MGC:41579
IMAGE:1246740, mRNA, complete cds.
ACCESSION BC028657
VERSION BC028657.1 GI:20307055
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 581)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Marcello Bento Soares, Ph.D.
CDNA Library Preparation: Soares Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
<http://www.systemsbio.org>
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAK Plate: 66 Row: h Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency OK
analysis.
FEATURES
source location/Qualifiers
1..581
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="C57BL/6J"
/clone="MGC:41579 IMAGE:1246740"
/tissue_type="Mammary gland"
/clone_id="Soares_mammary_gland_NDMMG"
/lab_host="DH10B"
/note="Vector: pT73D-Pac"
6..380
/codon_start=1
/product="RIKEN cDNA 170008c22 gene"
/protein_id="AAH28657.1"
/db_xref="GI:20307055"
/db_xref="LocusID:73454"
/translation="MAADLQIPVDIDSDGVFVYLLRYHLEPFGDPAKCKEIVR
CYKMEYHADYIDKVSGLQNGYDCDCLGGRISSHSDRKIHVYGYMGYGRADHS
VSTERIKAKYIPDEVYTWADQY"
BASE COUNT 151 a 142 c 156 g 132 t
ORIGIN
Query Match 72.2%; Score 270.8; DB 10; Length 581;
Best Local Similarity 83.2%; Pred. No. 1.4e-51;
Matches 308; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 6 GGTGGGAGACCTCGCTCATCTGATGTGACATGATCGAGGGGCTCTCAAGTA 65
DB 8 GCGGCTACCTCGGTGATCTGATGTGACATGATCGAGGGGCTCTCAAGTA 65
QY 66 TGTGCTGATTCGAGTCCACTTACAGAGCCTTCTGCGGATCGGCTTCAAGTA 67
DB 68 TGTGCTGATTCGAGTCCACTTACAGAGCCTTCTGCGGATCGGCTTCAAGTA 125
QY 126 GATGTGGGGGCTACAGAGTGGCTGATGACATGCGGATCGGCTTCAAGTA 127
DB 128 AATGCTCGGTGCTACAGAGTGGCTGATGACATGCGGATCGGCTTCAAGTA 187

QY 186 CGACATGCAAGAACAGAGCTCGACCTGATGATGCTGCGGGGGGCGCATCTCCACCA 245
DB 188 CGAGCTGCAAGAGATGGCTATGACTGTGATGCTTACAGGGGGGCGCATCTCCACCA 247
QY 246 GAGTCAGGACAGAAAGATTCACGCTGATGCTATTCATGCGCTTATGCTGCGCACCA 305
DB 248 GAGCCAGGACAGAAAGATTCACGCTGATGCTTATGCTTACGCTGCGCGCACCA 307
QY 306 CGCATTTCAACTGAGAAATCAAGCCAGTACCCGATCGAGTCCAGTCCGCTAA 365
DB 308 CTCTGTTTCAACTGAGAAATCAAGCCAGTATCCTGACTATGAGTACCTGCGCTGA 367
QY 366 CGAGCGCTAC 375
DB 368 CGATGCTAC 377
RESULT 10
AX305959
LOCUS AX305959
DEFINITION Sequence 710 from Patent WO018188.
ACCESSION AX305959
VERSION AX305959.1 GI:17645316
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 710)
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 018188-A 710 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source location/Qualifiers
1..975
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 199 a 176 c 369 g 231 t
ORIGIN
Query Match 63.3%; Score 237.4; DB 6; Length 975;
Best Local Similarity 80.0%; Pred. No. 5.2e-44;
Matches 292; Conservative 0; Mismatches 71; Indels 2; Gaps 1;
QY 6 GGTGGGAGACCTCGCTCATCTGATGTGACATGATCGAGGGGCTCTCAAGTA 65
DB 81 GCGGCTACCTCGGTGATCTGATGTGACATGATCGAGGGGCTCTCAAGTA 140
QY 66 TGTGCTGATTCGAGTCCACTTACAGAGCCTTCTGCGGATCGGCTTCAAGTA 140
DB 141 TGTGCTGATTCGAGTCCACTTACAGAGCCTTCTGCGGATCGGCTTCAAGTA 125
QY 126 GATGTGGGGGCTACAGAGTGGCTGATGACATGCGGATCGGCTTCAAGTA 200
DB 201 AATGCTCGGTGCTACAGAGTGGCTGATGACATGCGGATCGGCTTCAAGTA 185
QY 186 CGACATGCAAGAACAGAGCTCGACCTGATGATGCTGCGGGGGGCGCATCTCCACCA 245
DB 261 CGAGCTGCAAGAGATGGCTATGACTGTGATGCTTACGCTGCGCGCACCA 320
QY 246 GAGTCAGGACAGAAAGATTCACGCTGATGCTATTCATGCGCTTATGCTGCGCACCA 305
DB 321 GAGCCAGGACAGAAAGATTCACGCTGATGCTTATGCTTACGCTGCGCGCACCA 380
QY 306 CGCATTTCAACTGAGAAATCAAGCCAGTACCCGATCGAGTCCAGTCCGCTAA 365
DB 381 CTCTGTTTCAACTGAGAAATCAAGCCAGTATCCTGACTATGAGTACCTGCGCTGA 438
QY 366 CGAGC 370
DB 439 CGAGC 443

[illegible]

RESULT 12	369 bp	DNA	linear	PAT 14-JUN-2002
LOCUS AX410864/C	3511 from Patent WO0229103.			
DEFINITION	Sequence 3511 from Patent WO0229103.			
ACCESSION	AX410864			
VERSION	AX410864.1			
KEYWORDS	GI:21443569			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
JOURNAL	Gene expression profiles in liver cancer			
FEATURES	Patent: WO 0229103-A 3511 11-APR-2002;			
source	GENE LOGIC INC (US)			
location/Qualifiers				
1. 369				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/note="EMBL/Genbank Accession No. W42627"				
67 a 97 c 116 g 80 t 9 others				
BASE COUNT	67 a 97 c 116 g 80 t 9 others			
ORIGIN				
Query Match	38.7%; Score 145; DB 6; Length 369;			
Best Local Similarity	98.8%; Pred. No. 6, 1e-23;			
Matches 167; Conservative	0; Mismatches 0; Indels 2; Gaps 2;			
OY 209	ACTGTGACGTCTCGGGCGGCGGCCCATCTCC-ACCAAGTCCAGCAAGAATTCAC 267			
DB 346	ACTGGAATGTCTGGGGGGGCGCATCTCCCAACAGATCAGGCAAGAAATTCAC 287			
OY 268	GTGTACGGCTATTCCTCA-TGGCCATGTGCTCTCCCGACGACGCCATTTCAACTGAAAT 326			
DB 286	GTGTACGGCTATTCCTCAANTGGCTATGTCTGCTGCCAGCAGCGCATTTCAACTGAAAT 227			
OY 327	CAAGGCAAGTACCCCGACTACGAGGTCACTCGGGGCTACGACGGCTAC 375			
DB 226	CAAGGCAAGTACCCCGACTACGAGGTCACTCGGGGCTACGACGGCTAC 178			
RESULT 13	123559 bp	DNA	linear	HTG 10-MAR-2002
AC091284/C	WORKING DRAFT SEQUENCE, 11 unordered			
LOCUS	AC091284			
DEFINITION	Mus musculus clone RP23-20612, WORKING DRAFT SEQUENCE, 11 unordered			
ACCESSION	AC091284			
VERSION	AC091284.4			
KEYWORDS	HTGS; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE	Mus musculus.			
ORGANISM	Mus musculus.			
REFERENCE	1. (bases 1 to 123559)			
AUTHORS	Birten, B., Linton, L., Nusbaum, C. and Lander, E.			
TITLE	Mus musculus, clone RP23-20612			
JOURNAL	Unpublished			
AUTHORS	2. (bases 1 to 123559)			
TITLE	Birten, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,			
AUTHORS	Barra, N., Basile, J., Boguslavsky, I., Boukhalter, B., Brown, A.,			
TITLE	Camrata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,			
AUTHORS	Collins, S., Collimore, A., Cooke, P., Dearellano, K., Dewar, K.,			
TITLE	Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D.,			
AUTHORS	Galagan, J., Gardyna, S., Glinde, S., Goyette, M., Graham, L.,			
TITLE	Grand-pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, M.,			
AUTHORS	Iliev, I., Johnson, R., Jones, C., Karatas, A., Labouque, K.,			
TITLE	Lamares, R., Landers, T., Lebeck, J., Levine, R., Liu, G.,			
AUTHORS	Maclean, C., Macdonald, P., McQuinn, N., Matthews, C., Menes, L.,			
TITLE	McEwan, P., McKernan, K., McPherson, R., Melidrum, C., Norbu, C.,			
AUTHORS	Mihova, T., Minge, V., Murphy, T., Naylor, J., Nguyen, C., Oliver, J.,			
TITLE	Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,			
AUTHORS	Peterson, K., Phunhahng, P., Pierre, N., Pollara, V., Raymond, C.,			
TITLE	Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,			

Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuppach, R., Seaman, S., Sever, P., Sougniez, C., Spencer, B., Stange-Thomas, N., Stojanovic, N., Strauss, N., Subramanian, A., Talman, J., Tsafaye, S., Theodorou, J., Travers, M., Trivelpiece, N., Triggiani, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, D., Zemek, L., Zimmer, A., and Zody, M.

Submitted (12-APR-2007) | White, A.

3 (bases 1 to 123559)
Birren, B., Linton, J., Nuebaum, C.
Center for Genome
Cambridge, MA 02141, USA

TITLE
JOURNAL

COMMENT

On Mar 10, 2020 this sequence version replaced g1:15028547.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu>

Center, white-tailed eagle, www.genomecenter.ucsf.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.wi.mit.edu>

Contact: <http://www-seq.wi.mit.edu>
sequence submissions@seq.wi.mit.edu

Project Information
Center project name: T13330

Center clone name: 206_L_2

Summary Statistics
vector: plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly: 100% of reads

Assembly program: Phrap; version 0.960731
Consensus quality: 119904

Consensus quality:	bases at least Q40
119504	121642

Insert size: 200000: agarose gel

```
Insert size: 122559; sum-of-contigs
```

Quality coverage: 31.4 in Q20 bases; agarose-fp

sum-of-contrasts in Q20 bases; sum-of-contrasts

... consists of 11 conflice ... sequence. It currentl...

not known and their order in this sequence

... sequence record 1
... contigs are represented as

record will be updated with the exact sizes of the gaps are unknown

soon as it is available and the accession number will be preserved.

758: cont'd of 758 b

Accession	Gap of	bp in length
759	858	100
859	1603	100

1693 1792: contlig of 834 bp in length
1792: gap of 100 bp

1/93 6191: contig of 4399 bp in Jentat
6192 6301: 21

100 bp

FEATURES	6292	10475: contlg of 4184 bp in length
*	10476	10575: gap of 100 bp
*	10576	18552: contlg of 7977 bp in length
*	18553	18652: gap of 100 bp
*	18653	31018: contlg of 13366 bp in length
*	31019	31118: gap of 100 bp
*	31119	40891: contlg of 9773 bp in length
*	40892	40991: gap of 100 bp
*	40992	59807: contlg of 18816 bp in length
*	59808	59907: gap of 100 bp
*	73279	73278: contlg of 13371 bp in length
*	73379	73378: gap of 100 bp
*	102865	102865: contlg of 29487 bp in length
*	102866	102865: gap of 100 bp
*	102966	123599: contlg of 20554 bp in length
	Location	70,3

```

misc_feature      1.7568      nrc123 female Mouse BAC"
misc_feature      /note="assembly-fragment"
misc_feature      859.1692
misc_feature      /note="assembly-fragment"
misc_feature      1793.6191
misc_feature      /note="assembly-fragment"
misc_feature      6292.10475
misc_feature      /note="assembly-fragment"
misc_feature      10576.18552
misc_feature      /note="assembly-fragment"
misc_feature      18653.31018
misc_feature      /note="assembly-fragment"
misc_feature      31119.40891
misc_feature      /note="assembly-fragment"
misc_feature      40992.59807
misc_feature      /note="assembly-fragment"
misc_feature      59908.77278
misc_feature      /note="assembly-fragment"
misc_feature      73379.102865
misc_feature      /note="assembly-fragment"
misc_feature      102966.123559
misc_feature      /note="assembly-fragment"
BASE COUNT      29054 a 33278 c 32519 g 27698 t 1010 others
ORIGIN

```

[illegible]

RESULT 14	AL732590/c			
LOCUS	AL732590	236240 bp	DNA	linear HTG 25-JUN-2002
DEFINITION	Mus musculus chromosome 2 clone RP23-464C2, *** SEQUENCING IN			
PROGRESS	***, 3 unordered pieces.			
ACCESSION	AL732590			
VERSION	AL732590.4	GI:21615736		
KEYWORDS	HTG; HTGS-PHASE1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			

us-09-914-831-1.rge

OY	126	GATCGTCGGGCTACAAAGTGTGCATTGCCAATCAACAAGAATCNC	182
Dbl	106955	AATCGTGCGTACTAACAAATGGGCTGAATACCAGGTGAGGGCCGGGCTCTGTGCC	106899
RESULT 15			
AC103205		101816 bp DNA linear HTG 13-JUL-2007	
LOCUS		Rattus norvegicus clone CHZ30-32r23, *** SEQUENCING IN PROGRESS	
: DEFINITION		***, 56 unordered pieces.	
AC103205			
AC103205			
ACCESSION		GI:21731166	
VERSION		AC103205.3	
KEYWORDS		HTG; HTGS_PHASE1.	
		Norway rat.	

[illegible]

Query Match	Best Local Similarity	Matches 118;	Conservative	Score 98.2;	DB 2;	Length 101816;
32719	34339:	contig of 1621 bp in length				
34340	34439:	gap of unknown length				
34440	36416:	contig of 1977 bp in length				
36417	36516:	gap of unknown length				
36517	37749:	contig of 1233 bp in length				
37750	37849:	gap of unknown length				
37850	39344:	contig of 1495 bp in length				
39345	40929:	contig of 1485 bp in length				
39445	41029:	gap of unknown length				
40930	42630:	contig of 1600 bp in length				
41030	42729:	gap of unknown length				
42630	43775:	contig of 1045 bp in length				
43775	43874:	gap of unknown length				
43875	44956:	contig of 1082 bp in length				
44957	45056:	gap of unknown length				
45057	46229:	contig of 1172 bp in length				
46229	46329:	gap of unknown length				
46329	47673:	contig of 1345 bp in length				
47674	47773:	gap of unknown length				
47774	49401:	contig of 1628 bp in length				
49402	49501:	gap of unknown length				
49502	51309:	contig of 1808 bp in length				
51310	51409:	gap of unknown length				
51410	52832:	contig of 1423 bp in length				
52833	52932:	gap of unknown length				
52933	53435:	contig of 1413 bp in length				
53436	54445:	gap of unknown length				
54446	56298:	contig of 1853 bp in length				
56299	56398:	gap of unknown length				
56399	58622:	contig of 2224 bp in length				
58623	58722:	gap of unknown length				
58723	60492:	contig of 1770 bp in length				
60493	60592:	gap of unknown length				
60593	62311:	contig of 1719 bp in length				
62312	62412:	gap of unknown length				
62413	64048:	contig of 1636 bp in length				
64048	65147:	gap of unknown length				
65148	65957:	contig of 1809 bp in length				
65957	66036:	gap of unknown length				
66037	68350:	contig of 2294 bp in length				
68351	68450:	gap of unknown length				
68451	69892:	contig of 1442 bp in length				
69893	71854:	contig of 1862 bp in length				
71855	71955:	gap of unknown length				
71956	74448:	contig of 2494 bp in length				
74449	74549:	gap of unknown length				
74549	76920:	contig of 2372 bp in length				
76921	77020:	gap of unknown length				
77021	79592:	contig of 2572 bp in length				
79593	81310:	gap of unknown length				
81311	81411:	contig of 1618 bp in length				
81411	84051:	gap of unknown length				
84052	84151:	gap of unknown length				
84152	86734:	contig of 2573 bp in length				
86735	86834:	gap of unknown length				
86835	91125:	contig of 4291 bp in length				
91126	91225:	gap of unknown length				
91226	93747:	contig of 2322 bp in length				

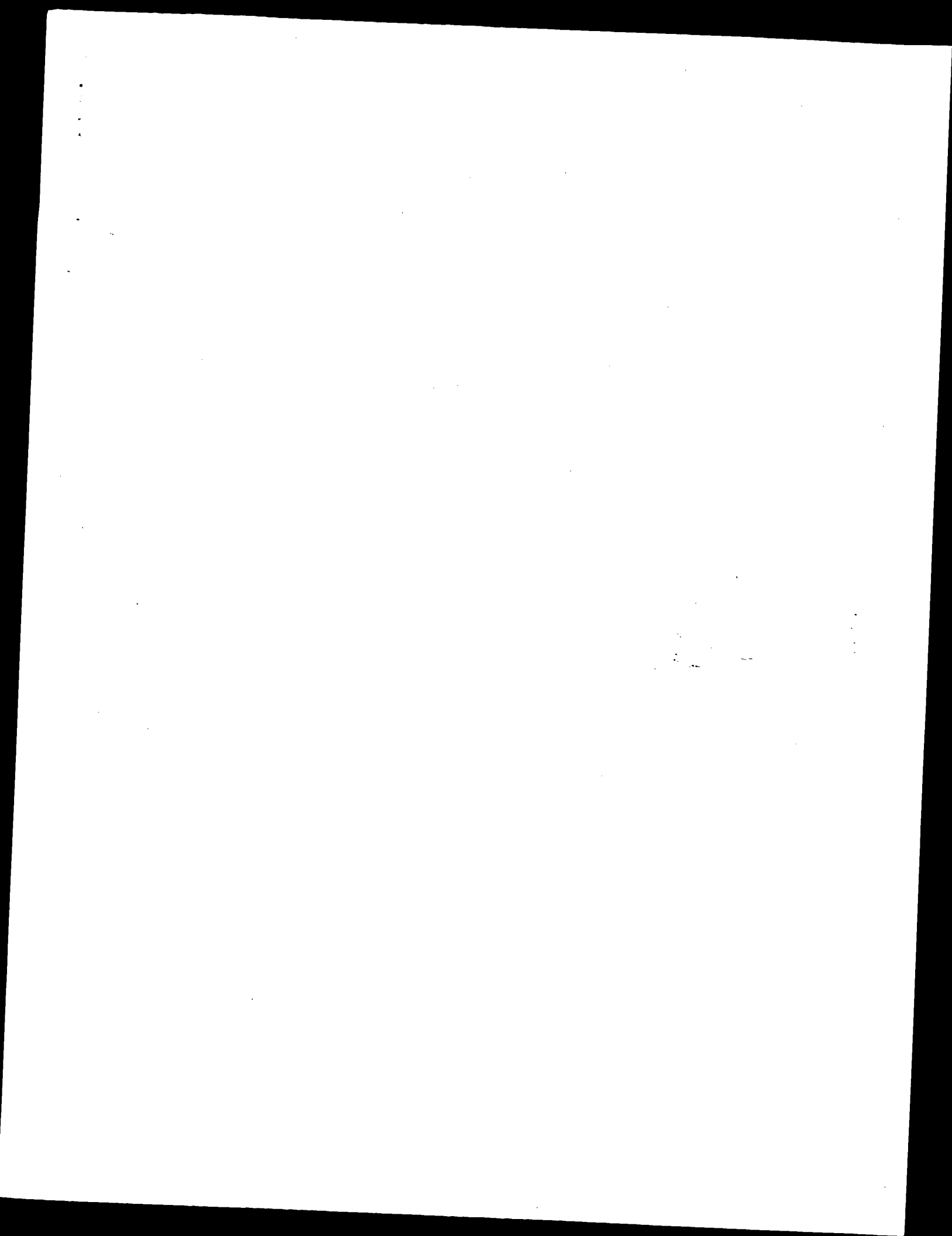
Wed May 14 09:22:02 2003

us-09-914-831-1.rge

Page 11

OY 280 TCATGCGCTATGCTCCTGCCGACAGCCCA 310
|||
|||
Db 82010 TCTATGCTGAGCTTCAGTCCCGGCTCTCCA 82040
|||
|||

Search completed: May 13, 2003, 19:00:08
Job time : 2304 secs



Title: 05-09-914-001 375

Sequence:

Scoring table:

IDENT142=non
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

post-processing:  Minimum Match 0%
                  Maximum Match 100%

```

Listing first 45 submitters

```
Database : N_Geneseq_101002
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
```

1: N_geneseq_010002:*

2: /SID52/gcgdata/genseq/genseqn_emb1/NA1960.DAT.*

3: /SID52/gcgdata/genseq/genseqn_emb1/NA1961.DAT.*

4: /SID52/gcgdata/genseq/genseqn_emb1/NA1962.DAT.*

5: /SID52/gcgdata/genseq/genseqn_emb1/NA1963.DAT.*

6: /SID52/gcgdata/genseq/genseqn_emb1/NA1964.DAT.*

7: /SID52/gcgdata/genseq/genseqn_emb1/NA1965.DAT.*

8: /SID52/gcgdata/genseq/genseqn_emb1/NA1966.DAT.*

9: /SID52/gcgdata/genseq/genseqn_emb1/NA1967.DAT.*

10: /SID52/gcgdata/genseq/genseqn_emb1/NA1968.DAT.*

11: /SID52/gcgdata/genseq/genseqn_emb1/NA1969.DAT.*

12: /SID52/gcgdata/genseq/genseqn_emb1/NA1970.DAT.*

13: /SID52/gcgdata/genseq/genseqn_emb1/NA1992.DAT.*

14: /SID52/gcgdata/genseq/genseqn_emb1/NA1993.DAT.*

15: /SID52/gcgdata/genseq/genseqn_emb1/NA1994.DAT.*

16: /SID52/gcgdata/genseq/genseqn_emb1/NA1995.DAT.*

17: /SID52/gcgdata/genseq/genseqn_emb1/NA1996.DAT.*

18: /SID52/gcgdata/genseq/genseqn_emb1/NA1998.DAT.*

19: /SID52/gcgdata/genseq/genseqn_emb1/NA1999.DAT.*

20: /SID52/gcgdata/genseq/genseqn_emb1/NA2000.DAT.*

21: /SID52/gcgdata/genseq/genseqn_emb1/NA2001.DAT.*

22: /SID52/gcgdata/genseq/genseqn_emb1/NA2002.DAT.*

23: /SID52/gcgdata/genseq/genseqn_emb1/NA2002.DAT.*

24: /SID52/gcgdata/genseq/genseqn_emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	375	100.0	375	21	AAA64502	DNA encoding a hum
2	375	100.0	1001	20	AA233497	Human prostate can
3	375	100.0	1129	22	AA160603	Human polynucleoti
4	375	100.0	1274	22	AA158817	Human polyomucleoti
5	375	100.0	4949	21	AA766889	Human ORF2544
6	296.8	79.1	487	23	ABY56855	Human prostate exp
7	296	76.3	629	21	AACT0901	Human secreted pro
8	286	78.9	1338	21	AACT6585	Human ORF2540
9	237.4	63.3	975	24	AB199674	Mouse ischaemic co

C	10	145	38.7	369	24	ANB97013
C	11	122.6	32.2	497	23	ANB45780
C	12	120.6	32.2	453	23	ABV15981
C	13	89.4	23.8	380	24	ANB96369
C	14	64	17.1	387	22	ANB65305
C	15	64	17.1	395	22	ANB65604
C	16	60.8	16.2	655	25	ABU103695
C	17	60	16.0	60	24	ABN43332
C	18	49.4	13.2	483	23	ABN120659
C	19	42.2	11.3	411	22	ABH87827
C	20	42.2	11.3	609	23	ABU103887
C	21	39.6	10.6	2544	23	AA556068
C	22	39.4	10.5	2569	23	ABU120692
C	23	38.8	10.5	8807	23	ABU12938
C	24	38.8	10.3	4407	23	ABU12349
C	25	38.8	10.3	6561	23	ABU12348
C	26	37.4	10.0	2436	23	AA552970
C	27	37.4	10.0	2636	23	ABU103906
C	28	37.4	10.0	2819	23	ABU103894
C	29	36.8	9.8	3205	23	ABU103630
C	30	36.8	9.8	1989	21	AA664030
C	31	36.8	9.7	61557	23	AA559521
C	32	36.4	9.7	6164	22	AAH98413
C	33	35.6	9.5	396	13	AAH98413
C	34	35.6	9.5	396	14	AAQ21833
C	35	35.6	9.5	396	22	AA368559
C	36	35.6	9.5	390	24	AAQ76510
C	37	35.6	9.5	17575	23	AAI72175
C	38	35.2	9.4	15370	23	ABQ888207
C	39	35.2	9.4	480	22	ABQ50907
C	40	34.6	9.2	1072	15	AAQ65408
C	41	34.2	9.1	1693	24	ABR34442
C	42	34.2	9.1	870	22	AAH03832
C	43	34.2	9.1	2012	22	AAH53563
C	44	34.2	9.1	2370	24	ABQ84808
C	45	34.2	9.1	2380	21	AAAC768535
C	46	34.2	9.1	1062	23	ABU434367

ALIGNMENTS

Gene#511 used to
Human prostate exp
Human prostate exp
to gene #487 used to
Novel human polyu
Novel human polyu
Drosophila melanog
Human spliced tran
Drosophila melanog
Pepperwint plant o
Drosophila melanog
Salmonella typhi D
Drosophila melanog
Drosophila melanog
Drosophila melanog
Drosophila melanog
Klebsiella pneumon
Drosophila melanog
Drosophila melanog
Drosophila melanog
Winter wheat inver
Prophionibacterium
Human EST-derived
Human EST-derived
Randomising oligon
PCR primer for 5'
Sequence containin
Oligo #7 for cloni
Human osteoblast d
Drosophila melanog
cDNA encoding join
Human breast speci
Human cDNA clone (c
Human cDNA sequenc
Human osteoblast d
Human OREX ORE413
Drosophila melanog

XX	RESULT 1	
XX	AAA64502	
XX	ID	AAA64502 standard; DNA; 375 BP.
XX	AC	
XX	AAA64502;	
XX	02-JAN-2001	(first entry)
XX	DNA encoding a human histidine protein phosphatase polypeptide.	
XX	Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;	
XX	cell regulation; cell growth; cancer; immune disorder; viral infection	
XX	genetic disorder; heart disease; N-phosphorylation imbalance; ss.	
XX	Homo sapiens.	
XX		
OS		
XX	Key	Location/Qualifiers
XX	Key	1.375
XX	CDS	/tag= a
XX		/product= "histidine protein phosphatase"
XX		/note= "no termination codon given"
XX		
XX	WO200052175-A1	
XX		
XX	08-SEP-2000.	
XX		
XX	02-MAR-2000; 2000MO-EP01774.	
XX		
XX	04-MAR-1999; 99DE-1009388.	
XX		
XX	(MERE) MERCK PATENT GMBH.	
XX		

PI Klump S, Kellner R;
 XX WPI: 2000-572187/53.
 DR P-PSDB; AAB08787.
 XX

PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
 PT immune disorders, viral infection, genetic disorders, and heart disease
 XX
 PS Claim 9; Page 18; 39pp; English.

CC The present sequence encodes a human histidine phosphatase polypeptide.
 CC The polypeptide has a high specificity for phosphohistidine and a
 CC molecular weight of 1300-1500 Da. The histidine phosphatase gene is
 CC localized on chromosome 9 (9q33). The histidine phosphatase can be used
 CC for diagnosis and treatment of pathological states of cell regulation
 CC and cell growth. These include cancers, immune disorders, viral
 CC infection, genetic disorders, and heart disease. The histidine
 CC phosphatase can also be used for identifying agonists and antagonists
 CC which can be used to treat conditions associated with N-phosphorylation
 CC imbalance.

SO Sequence 375 BP; 83 A; 110 C; 112 G; 70 T; 0 other;
 Query Match 100.0%; Score 375; DB 21; Length 375;
 Best Local Similarity 100.0%; Pred. No. 9e-97; 0; Mismatches 0; Indels 0; Gaps 0;

Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGAGACATCGACGCGCTCTTC 60
 DB 1 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGAGACATCGACGCGCTCTTC 60
 QY 61 AAGTATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 60
 DB 61 AAGTATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 60
 QY 121 AAGGATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 120
 DB 121 AAGGATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 120
 QY 181 TCGGGCGACATGCAAGCAAGGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 180
 DB 181 TCGGGCGACATGCAAGCAAGGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 180
 QY 241 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 240
 DB 241 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 240
 QY 301 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 300
 DB 301 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 300
 QY 361 GCTAAGCAGCGGCTAC 375
 DB 361 GCTAAGCAGCGGCTAC 375

RESULT 2
 AAZ33497
 ID AAZ33497 standard; cDNA; 1001 BP.
 AC AAZ33497;
 XX
 DT 08-DEC-1999 (first entry)
 XX
 DE Human prostate cancer-associated EST 21.
 XX
 KW Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
 XX
 OS cancer; tissue specificity; human; ss.
 XX
 PN Homo sapiens.
 DE19811194-A1.

XX
 PD 16-SEP-1999.
 XX
 PF 10-MAR-1998; 98DE-1011194.
 XX
 PR 10-MAR-1998; 98DE-1011194.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX
 DR WPI: 1999-519629/44.
 DR P-PSDB; AAY48347, AAY48348.

PT New nucleic acid expressed at high level in normal prostatic tissue and
 PT encoded polypeptides, used to treat cancer and screen for therapeutic
 PS agents
 PS Claim 1a; 92; 194pp; German.

CC This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)
 CC encoded by (A) are used: (a) for identifying agents for treatment of
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally
 CC where expressed by gene therapy methods. (A) is also used to isolate
 CC full-length genes (for gene therapy) and for recombinant production of
 CC by assembly of ESTs (expressed sequence tags) before these are analyzed
 CC for expression pattern (tissue specificity). This approach eliminates
 CC many of the false results, as regards tissue specificity, associated
 CC with known methods that use single (usually short) ESTs. AAZ33477-233540
 CC represent expressed sequence tags described in the method of the
 CC invention.

SO Sequence 1001 BP; 178 A; 312 C; 329 G; 182 T; 0 other;
 Query Match 100.0%; Score 375; DB 20; Length 1001;
 Best Local Similarity 100.0%; Pred. No. 1.2e-96;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGAGACATCGACGCGCTCTTC 60
 DB 435 ATGGCGGTGGCGGACCTGCTCTCATTCCTGATGAGACATCGACGCGCTCTTC 60
 QY 61 AAGTATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 60
 DB 495 AAGTATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 60
 QY 121 AAGGATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 120
 DB 555 AAGGATGTCGTGATCCGAGTCCGCTCCGCTCCGCGGCGCGGCGCGCTCCGCTCC 120
 QY 181 TCGGGCGACATGCAAGCAAGGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 180
 DB 615 TCGGGCGACATGCAAGCAAGGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 180
 QY 241 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 240
 DB 675 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 240
 QY 301 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 300
 DB 735 CACGAGATCGTGGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 300
 QY 361 GCTAAGCAGCGGCTAC 375
 DB 795 GCTAAGCAGCGGCTAC 809

RESULT 3
 AA160603/C
 ID AA160603 standard; cDNA; 1129 BP.
 XX

Db 547 CACCAGAGTCAGACAAAGATGACGTGATGACGCTATTCATGCGCTATGCTGCTG 606
 Oy 301 CAGCAGCCATTTCACTGAGAAATCAAAAGCAAGTACCCGACTACGAGTCCCTG 360
 Db 607 CAGCAGCCATTTCACTGAGAAATCAAAAGCAAGTACCCGACTACGAGTCCCTG 666
 Oy 361 GCTAAGCAGCGCTAC 375
 Db 667 GCTAAGCAGCGCTAC 681
 RESULT 6
 ABV56855 standard; cDNA: 487 BP.
 ID ABV56855:
 AC ABV56855:
 XX 17-SEP-2002 (first entry)
 DT Human prostate expression marker cDNA 56846.
 DE Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX WO200160860-A2.
 PN 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US05171.
 PF 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 PI WPI: 2001-662795/76.
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1: Page 10954; 11750pp: English.
 PS The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
 CC Sequence 487 BP; 112 A; 144 C; 131 G; 100 T; 0 other;
 SQ Query Match 79.1%; Score 296.8; DB 23; Length 487;
 Best Local Similarity 88.6%; Pred. No. 1,5e-74;
 Matches 333; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

Oy 1 ATGGCGTGGGGGAGCCTGCTCTCATTCGATGAGACATCG-ACGCGAGCGCTCTT 59
 Db 47 ATGGCGGGGGGAGCCTGCTCTCATTCGATGAGACATCG-ACGCGAGCGCTCTT 106
 Oy 60 CAAGTATGCTGATTCGGAATCCAGTCCGCTCCGCGGGGCTCCGCTGCAAGAG 119
 Db 107 AAAGTTGTCGTATTCGGAATCCAGTCCGCTCCGCGGGGCTCCGCTGCAAGAG 166
 Oy 120 CAAGGAGATGCTGCGGGCTACAGTGGGCTGATGATGAGGAGCATATCAACAAAGT 179
 Db 167 CAAGGAGATGCTGCGGGCTACAGTGGGCTGATGATGAGGAGCATATCAACAAAGT 226
 Oy 180 GTGCGGGGAGCATGCAAGCAAGGCTGAGTGTGCTGCGGGGCGGCGGCAATTC 239
 Db 227 GTGCGGGGAGCATGCAAGCAAGGCTGAGTGTGCTGCGGGGCGGCGGCAATTC 286
 Oy 240 CCACGAGATGCAAGCAAGGCTGAGTGTGCTGCGGGGCGGCGGCAATTC 299
 Db 287 CCACGAGATGCAAGCAAGGCTGAGTGTGCTGCGGGGCGGCGGCAATTC 346
 Oy 300 CCAGCAGCGCTATTCATGAGAAATCAAAAGCAAGTACCCGACTACGAGTCCCTG 359
 Db 347 CCAGCAGCGCTATTCATGAGAAATCAAAAGCAAGTACCCGACTACGAGTCCCTG 406
 Oy 360 GGCTAACGAGCGGCTAC 375
 Db 407 GGCTAACGAGCGGCTAC 422
 RESULT 7
 AAC10901 standard; cDNA: 629 BP.
 ID AAC10901:
 AC AAC10901:
 XX 06-OCT-2000 (first entry)
 DT Human secreted protein 5' EST, SEQ ID NO: 14976.
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 OS Homo sapiens.
 XX EP1033401-A2.
 PN 06-SEP-2000.
 PD 21-FEB-2000; 2000EP-0200610.
 PR 26-FEB-1999; 99US-0122487.
 XX (GIST) GENSET.
 PA Dumas Milne Edwards J, Duclert A, Giordano J;
 PI WPI: 2000-500381/45.
 DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1: SEQ ID 14976; 71pp + CD-ROM: English.
 PS The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or poly(A⁺) RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer

CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

Sequence 629 BP; 135 A; 180 C; 193 G; 115 T; 6 other;

Query Match 78.9%; Score 296; DB 21; Length 629;
 Best Local Similarity 93.6%; Pred. No. 2,7e-74;
 Matches 320; Conservative 5; Mismatches 0; Indels 17; Gaps 1;

QY 1 ATGGCGTGGCGGACCTGCTCTATTCCTGATGTGACATCAGACGCGCTTC 60
 DB 305 ATGGCGTGGCGGACCTGCTCTATTCCTGATGTGACATCAGACGCGCTTC 60
 QY 61 AAGTATGCTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 364
 DB 365 AAGTATGCTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 120
 QY 121 AAGGAGATCGTGGCGGCTGACATCCGATCCGATCCGATCCGATCCGATCC 424
 DB 425 AAGGAGATCGTGGCGGCTGACATCCGATCCGATCCGATCCGATCCGATCC 180
 QY 181 TCGGCGACATCGACAGCAAGGCTGCTGATGTGACATCCGATCCGATCCGATCC 467
 DB 468 TCGGCGACATCGACAGCAAGGCTGCTGATGTGACATCCGATCCGATCCGATCC 240
 QY 241 CACCAAGTGCAGCAAGCAAGATTCAGTTCGATCCGATCCGATCCGATCCGATCC 527
 DB 528 CACCAAGTGCAGCAAGCAAGATTCAGTTCGATCCGATCCGATCCGATCCGATCC 300
 QY 301 CACCAAGTGCAGCAAGCAAGATTCAGTTCGATCCGATCCGATCCGATCCGATCC 587
 DB 588 CACCAAGTGCAGCAAGCAAGATTCAGTTCGATCCGATCCGATCCGATCCGATCC 342
 ID AACT6985
 AC AACT6985 standard; cDNA; 1338 BP.

RESULT 8
 AACT6985
 ID AACT6985 standard; cDNA; 1338 BP.
 AC AACT6985:
 DT 08-FEB-2001 (first entry)
 XX Human ORF2540 polynucleotide sequence SEQ ID NO:5079.

Human: open reading frame; ORF; detection; cytosolic; hepatotropic;
 KW vulnereary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 KW anticonvulsant; osteopathic; antidiabetic; immunosuppressant; cardiant;
 KW hypotensive; dermatological; coagulant; vasotropic; antidiabetic;
 KW antiviral; antibacterial; immunosuppressive; antidiabetic;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW severe combined immunodeficiency; systemic lupus erythematosus; infection;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW thrombosis; contraceptive; antiinflammatory disease; coagulation;
 XX Homo sapiens.

OS Homo sapiens.
 XX MO200058473-A2.
 PD 05-OCT-2000.
 XX 31-MAR-2000; 2000WO-US08621.
 PE 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI: 2000-602362/57.
 DR P-PSDB; AAB42776.

PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PS Claim 5; Page 4263-4264; 5507pp; English.

AC74446 to AACT7606 encode the proteins given in AAB40237 to AAB43397,
 which represent the human ORF open reading frames 1 to 3161. The ORF
 sequences have activities such as: cytosolic; hepatotropic; vulnereary;
 CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 CC osteopathic; anticonvulsant; antidiabetic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; immunosuppressive;
 CC antidiabetic; hypotensive; dermatological; coagulant; vasotropic;
 CC antidiabetic; antianemic; antiviral; antifungal; antirheumatic;
 CC antihypertensive; antiparkinsonian; antidiabetic; antidiabetic;
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder.
 CC The presence of or predisposition to, or preventing or treating
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterolemia, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

Sequence 1338 BP; 249 A; 455 C; 396 G; 238 T; 0 other;

Query Match 76.3%; Score 286; DB 21; Length 1338;
 Best Local Similarity 100.0%; Pred. No. 2.4e-71;
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTGGCGGACCTGCTCTATTCCTGATGTGACATCAGACGCGCTTC 60
 DB 340 ATGGCGTGGCGGACCTGCTCTATTCCTGATGTGACATCAGACGCGCTTC 60
 QY 61 AAGTATGCTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 399
 DB 400 AAGTATGCTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 120
 QY 121 AAGGAGATCGTGGCGGCTGACATCCGATCCGATCCGATCCGATCCGATCC 459
 DB 460 AAGGAGATCGTGGCGGCTGACATCCGATCCGATCCGATCCGATCCGATCC 180
 QY 181 TCGGCGACATCGACAGCAAGGCTGCTGATGTGACATCCGATCCGATCCGATCC 519
 DB 520 TCGGCGACATCGACAGCAAGGCTGCTGATGTGACATCCGATCCGATCCGATCC 240
 QY 241 CACCAAGTGCAGCAAGCAAGATTCAGTTCGATCCGATCCGATCCGATCC 579
 DB 580 CACCAAGTGCAGCAAGCAAGATTCAGTTCGATCCGATCCGATCCGATCC 286
 ID ABI99674
 AC ABI99674 standard; cDNA; 975 BP.

RESULT 9
 ABI99674
 ID ABI99674 standard; cDNA; 975 BP.
 AC ABI99674:
 DT 07-MAR-2002 (first entry)

Db

226

CAAAAGCCAGTACCGGCTACGAGTGTGCGGTAAACGCGGTAC 178

RESULT 11

ID

ABV45780 standard; cDNA, 497 BP.

AC

ABV45780;

16-SEP-2002 (first entry)

Human prostate expression marker cDNA 45771.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

20-FEB-2001; 2001WO-US05171.

17-FEB-2000; 2000US-183319P.

16-MAR-2000; 2000US-189862P.

25-MAY-2000; 2000US-207454P.

09-JUN-2000; 2000US-211314P.

18-JUL-2000; 2000US-219007P.

13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of

prostate cells and correlating with presence of prostate cancer, useful

for detecting presence of prostate cancer, stage of prostate cancer

Claim 1; Page 9050; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising

a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate

cancer in a patient;

(d) selecting a composition for inhibiting prostate cancer

(e) assessing the prostate cell carcinogenic potential of a compound;

(f) determining whether prostate cancer has metastasized in a patient;

(g) assessing the aggressiveness or indolence of prostate cancer in a

(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 497 BP; 97 A; 133 C; 142 G; 125 T; 0 other;

Query Match

Best Local Similarity 32.7%; Score 122.6; DB 23; Length 497;

Matches 123; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

153 GTACATGCGGACATCTAGACAAAGTGTGCGGACATGCAACAAAGGCTGCGACTG 212

Db 57 GTACATGCGGACATCTAGACAAAGTGTGCGGACATGCAACAAAGGCTGCGACTG 212

QY 213 TTAGTGTCTGGGCGGCGGCGCATCTCCACGAGATGAGCAAGAAGATTGACGCTA 272

Db 117 TTAGTGTCTGGGCGGCGGCGCATCTCCACGAGATGAGCAAGAAGATTGACGCTA 272

QY 273 CGGCTATTG 281

Db 177 CAGCCTCTC 185

RESULT 12

ID

ABV15981 standard; cDNA, 423 BP.

AC

ABV15981;

13-SEP-2002 (first entry)

Human prostate expression marker cDNA 15972.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

20-FEB-2001; 2001WO-US05171.

17-FEB-2000; 2000US-183319P.

16-MAR-2000; 2000US-189862P.

25-MAY-2000; 2000US-207454P.

09-JUN-2000; 2000US-211314P.

18-JUL-2000; 2000US-219007P.

13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of

prostate cells and correlating with presence of prostate cancer, useful

for detecting presence of prostate cancer, stage of prostate cancer

Claim 1; Page 2677; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising

a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate

cancer in a patient;

(d) selecting a composition for inhibiting prostate cancer

(e) assessing the prostate cell carcinogenic potential of a compound;

(f) determining whether prostate cancer has metastasized in a patient;

(g) assessing the aggressiveness or indolence of prostate cancer in a

(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 423 BP; 82 A; 111 C; 118 G; 112 T; 0 other;

Query Match

Best Local Similarity 32.2%; Score 120.6; DB 23; Length 423;

Matches 123; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

155 ACATCGGACATCTAGACAAAGTGTGCGGACATGCAAGAAGGCTGCGACTG 214

Db 1 ACCATCGGACATCTAGACAAAGTGTGCGGACATGCAAGAAGGCTGCGACTG 214

QY 215 AGTGTCTGGGCGGCGGCGCATCTCCACGAGATGAGCAAGAAGATTGACGCTA 274

DB 61 AGTGTCTGGGCGGCGGCGCATCTCCACGAGTACGAGCAAGATTACGCTTACA 120
 QY 275 GCTATTC 281
 11 11
 DB 121 GCCTCTC 127

RESULT 13
 ABN96369/C
 ID ABN96369 standard; DNA; 330 BP.

AC ABN96369;

DT 13-AUG-2002 (first entry)

DE Gene #2867 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumor; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN WO200229103-A2.

PD 11-APR-2002.

PF 02-OCT-2001; 2001WO-US30589.

PR 02-OCT-2000; 2000US-237054P.

PA (GENE-) GENE LOGIC INC.

PI Horne D, Alvarez C, Peres-Da-Silva S, Vockley JG;

DR WPI; 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer.

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample

PS Claim 1; SEQ ID NO 2867; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 330 BP; 57 A; 77 C; 114 G; 76 T; 6 other;

Query Match 23.8%; Score 89.4; DB 24; Length 330;

Best Local Similarity 95.7%; Pred. No. 9.5e-16; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 4;

QY 282 CATGGCTATGTCCTGCGCCAGACGCGCATTTCACTGAGAAATCAAGCCAGTACC 341

DB 285 CAAGGCTATGTCCTGCGCCAGACGCGCATTTCACTGAGAAATCAAGCCAGTACC 226

QY 342 CGACTAGAGGTCACTGGGCTAACGAGCGCTAC 375

DB 225 CGACTAGAGGTNACTCTGGGCTAACGAGCGCTAC 192

RESULT 14
 ID AAF66305
 AAF66305 standard; cDNA; 387 BP.

AC AAF66305;

DT 09-APR-2001 (first entry)

DE Novel human polynucleotide, SEQ ID NO: 2061.

XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 KW breast cancer; lung cancer; cancer detection; ss.

OS Homo sapiens.

PN WO200102568-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18374.

PR 02-JUL-1999; 99US-0142310.

PR 02-JUL-1999; 99US-0142311.

PA (CHIR) CHIRON CORP.

PA (HSE-) HSE INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;

PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;

PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;

PI Kita D, Garcia V, Jones LM, Strache-Crain B;

DR WPI; 2001-091805/10.

PT Library of polynucleotides for diagnosing a cancerous state of a

PT mammalian cell and detecting cancer, particularly of the colon or

PT prostate, comprises 3351 human polynucleotide sequences

PS Claim 9; Page 842; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.

SO Sequence 387 BP; 58 A; 145 C; 118 G; 66 T; 0 other;

Query Match 17.1%; Score 64; DB 22; Length 387;

Best Local Similarity 100.0%; Pred. No. 1.0e-08; Indels 0; Gaps 0;

Matches 64; Conservative 0; Mismatches 0;

QY 223 GCGCGCGCGCGCATCTCCACGAGTACGAGCAAGATTACGCTGATTC 282

DB 8 GCGCGCGCGCGCATCTCCACGAGTACGAGCAAGATTACGCTGATTC 67

QY 283 ATGG 286

DB 68 ATGG 71

Job time : 272 secs

```

RESULT 15
AAF66304 standard; cDNA; 395 BP.
ID AAF66304
AC AAF66304;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2060.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; ss.
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
FD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
XX 02-JUL-1999; 99US-0142311.
XX
PA (CHIR ) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamsom G, Drmanac R;
PI Ckenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences
XX
PS Claim 9; Page 842; 1046bp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping or
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 395 BP; 59 A; 145 C; 123 G; 68 T; 0 other;

```

Query Match 17.1%; Score 64; DB 22; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 223 GGGGGGGGGGATCCACAGAGTACAGAGATTCAGTACGGCTATGCC 282
DB 8 GGGGGGGGGGATCCACAGAGTACAGAGATTCAGTACGGCTATGCC 282
QY 283 ATGG 286
DB 68 ATGG 71

```

Search completed: May 13, 2003, 18:21:37

Wed May 14 09:22:04 2003

us-09-914-831-1.rst

Page 1

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2003, 17:39:57 ; Search time 1747 Seconds
(without alignments)
3476.421 Million cell updates/sec

Title: US-09-914-831-1

Perfect score: 375
Sequence: 1 atggcgagcgagcgagctgcg.....cttgagctaacgagcgctac 375

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_estchum:*
3: em_estlin:*
4: em_estmuv:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	375	100.0	452 9 AA461149	AA461149 zx64g11.r
2	375	100.0	545 13 B1753415	B1753415 603026309
3	375	100.0	561 10 AV693647	AV693647 AV693647
4	375	100.0	577 12 BG353346	BG353346 602404175
5	375	100.0	578 13 BM471874	BM471874 AGENCOURT
6	375	100.0	597 14 BM691037	BM691037 UI-R-C11-

7	375	100.0	599 14 BM705930	BM705930 UI-R-DW0-
8	375	100.0	604 13 B1833178	B1833178 603087211
9	375	100.0	607 12 BG819649	BG819649 602783132
10	375	100.0	612 14 BM918530	BM918530 AGENCOURT
11	375	100.0	866 13 B1522694	B1522694 603175733
12	375	100.0	1006 14 B0677546	B0677546 AGENCOURT
13	375	100.0	1010 12 BF345224	BF345224 602018170
14	373.4	99.6	579 12 BF405028	BF405028 UI-R-CA1-
15	373.4	99.6	589 14 B0651896	B0651896 AGENCOURT
16	373.4	99.6	617 13 BM550408	BM550408 AGENCOURT
17	373.4	99.6	621 14 B0678234	B0678234 AGENCOURT
18	373.4	99.6	657 13 B1912910	B1912910 603176712
19	373.4	99.6	668 10 AV686428	AV686428 AV686428
20	373.4	99.6	891 14 BM807502	BM807502 AGENCOURT
21	373.4	99.6	944 14 B0723150	B0723150 601463461
22	373.4	99.6	1081 12 BE777973	BE777973 603022468
23	373	99.5	605 13 B1752266	B1752266 603022468
24	373	99.5	606 13 B1838113	B1838113 603083612
25	373	99.5	830 12 BF342672	BF342672 602013620
26	372.4	99.2	588 14 B0017946	B0017946 UI-R-DH1-
27	372	99.2	588 13 B1522185	B1522185 603081118
28	372	99.2	583 13 B1521337	B1521337 603081118
29	372	99.2	601 12 BF341754	BF341754 602016245
30	372	99.2	616 14 BM919123	BM919123 AGENCOURT
31	371.8	99.1	616 9 AA554870	AA554870 n106104.8
32	371	98.9	858 13 B1754026	B1754026 603027468
33	371	98.9	880 14 B0683430	B0683430 AGENCOURT
34	370	98.7	627 12 BE901481	BE901481 601674940
35	368.8	98.3	560 14 W52042	W52042 zc47b1.r1
36	367.8	98.1	594 14 U83510	U83510 U83510 Soar
37	366	97.6	502 14 BM6944769	BM6944769 UI-R-C11-
38	365.4	97.4	651 10 AV704714	AV704714 AV704714
39	365	97.3	555 14 B0420364	B0420364 AGENCOURT
40	365	97.3	901 14 B0683740	B0683740 AGENCOURT
41	364	97.1	502 13 B1523429	B1523429 603175733
42	364	97.1	588 12 BF184250	BF184250 601843311
43	364	97.1	588 13 B1196173	B1196173 602754688
44	361.4	96.4	615 11 AF161490	AF161490 Homo sapi
45	359.4	95.8	862 12 BG036900	BG036900 602286821

ALIGNMENTS

RESULT 1
AA461149 452 bp mRNA linear EST 09-JUN-1997
zx64g11.r1 Soares total_fetus.Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:796292 5' similar to TR:G913978 G913978 JAN A. ; mRNA
sequence.

ACCESSION AA461149
VERSION AA461149.1 GI 2186269
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 452)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, D., Moore, B.,
Schellenberg, K., Stepien, M., Tan, F., Treising, B., White, Y., Wyllie,
T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham

TITLE JOURNAL
COMMENT

FEATURES High quality sequence stop: 404.
Location/Qualifiers
1. 452

/organism="Homo sapiens"
/db_xref="taxon:6040325"
/db_xref="taxon:9606"
/clone="IMAGE:796292"
/clone_11b="Soares_total_fetus_ND2HF8_9"
/dev_stage="8-9 weeks"
/lab_host="DH10B"

/note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand CDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer (5' TGTACCAATCTGAGTGGAGCGCCGCTTAATTTTATTTTATTTT 3'). Double-stranded CDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT3D vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 94 a 145 c 130 g 83 t
ORIGIN

Query Match Best Local Similarity 100.0%; Score 375; DB 9; Length 452;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCGGTGGCGGACCTCGCTCATTCCTGATGTGACATCGACTCGACGGGCTTTC 60
6 ATGGCGGTGGCGGACCTCGCTCATTCCTGATGTGACATCGACTCGACGGGCTTTC 65
61 AAGTATGTGCTATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 65
121 AAGGATATGTCGGGGCTACAAAGTGGGCTGAGTACATCGGACATCTACGAAAGTG 180
126 AAGGATATGTCGGGGCTACAAAGTGGGCTGAGTACATCGGACATCTACGAAAGTG 185
181 TCGGCGGACATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGG 240
186 TCGGCGGACATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGG 245
241 CACCAAGATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 300
246 CACCAAGATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 305
301 CAGCAGCATTCCTCACTGAGAAATCAAGCCAGTACCCGAGTACGAGGTCACCTGG 360
306 CAGCAGCATTCCTCACTGAGAAATCAAGCCAGTACCCGAGTACGAGGTCACCTGG 365
361 GCTAAGCAGCGCTAC 375
366 GCTAAGCAGCGCTAC 380

RESULT 2
LOCUS B1753415 545 bp mRNA linear EST 25-SEP-2001
DEFINITION 603026309F1 NIH_MGC_114 Homo sapiens CDNA clone IMAGE:5196579 5',
ACCESSION B1753415
VERSION B1753415
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC. http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cagabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11492 row: b column: 04
High quality sequence stop: 342.
Location/Qualifiers
1. 545

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196579"
/clone_11b="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCMV-SPORE6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH-MGC Library."

BASE COUNT 108 a 176 c 159 g 102 t
ORIGIN

Query Match Best Local Similarity 100.0%; Score 375; DB 13; Length 545;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGGCGGTGGCGGACCTCGCTCATTCCTGATGTGACATCGACTCGACGGGCTTTC 60
6 ATGGCGGTGGCGGACCTCGCTCATTCCTGATGTGACATCGACTCGACGGGCTTTC 65
61 AAGTATGTGCTATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 65
121 AAGGATATGTCGGGGCTACAAAGTGGGCTGAGTACATCGGACATCTACGAAAGTG 180
126 AAGGATATGTCGGGGCTACAAAGTGGGCTGAGTACATCGGACATCTACGAAAGTG 185
181 TCGGCGGACATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGG 240
186 TCGGCGGACATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGG 245
241 CACCAAGATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 300
246 CACCAAGATCGACAAAGAGGCTGAGTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 305
301 CAGCAGCATTCCTCACTGAGAAATCAAGCCAGTACCCGAGTACGAGGTCACCTGG 360
306 CAGCAGCATTCCTCACTGAGAAATCAAGCCAGTACCCGAGTACGAGGTCACCTGG 365
361 GCTAAGCAGCGCTAC 375
366 GCTAAGCAGCGCTAC 380

RESULT 3
LOCUS AV693647 561 bp mRNA linear EST 16-JAN-2002
DEFINITION AV693647 GKC Homo sapiens CDNA clone GKAMB03 5', mRNA sequence.
ACCESSION AV693647
VERSION AV693647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

Page 3

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
human.	Homo sapiens	1 (bases 1 to 577)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	NIH-MGC http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
				Email: cgabbs@email.nih.gov	Contact: Robert Strausberg, Ph.D.	
				Tissue procurement: ATCC	CDNA Library Preparation: Ling Hong/Rubin Laboratory	
				CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	DNA Sequencing by: Incyte Genomics, Inc.	
				Clone distribution: MGC clone distribution	Information can be	
				found through the I.M.A.G.E. Consortium/LLNL at:		
				http://image.llnl.gov		
				Plate: LICM1220	row: 6 column: 19	
				High quality	sequence stop: 576.	
FEATURES	Source	Location/Qualifiers				
		1..577				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/clone_xref="IMAGE:4541586"				
		/clone_lib="NIH MGC_21"				
		/issue_type="choriocarcinoma"				
		/lab_host="DH10B (phage-resistant)"				
		/note="Organ: Placenta; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming.				
		directionally cloned into EcoRI/XhoI sites using the				
		following 5' adaptor: GGGACGAG(G). Size-selected >500bp				
		for average insert size 1.8kb. Library constructed by				
		Ling Hong in the laboratory of Gerald M. Rubin (University				
		of California, Berkeley) using ZAP-CDNA synthesis kit				
		(Stratagene) and Superscript II RT (Life Technologies)."				
		(Stratagene) and Superscript II RT (Life Technologies)."				
		1 others				
BASE COUNT	130 a	179 c	163 g	104 t		
ORIGIN						
Query Match	100.0%;	Score 375;	DB 12;	Length 577;		
Best Local Similarity	130.0%;	Pred. No. 5,7e-86;				
Matches 375;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;	
1	ATGGCGGTGGCGGACCTCGCTCATCTTCGATGTGGACATCGACTCGAGGGCTCTTC	60				
7	ATGGGGGGGGGGACCTCGCTCATCTTCGATGTGGACATCGACTCGAGGGCTCTTC	66				
61	AAGTATGTGTCGATCGAGTCGACCTCGGCTCCGGCTCCGGGCTCGGGCTGCAGAGAGC	120				
67	AAGTATGTGTCGATCGAGTCGACCTCGGCTCCGGCTCCGGGCTCGGGCTGCAGAGAGC	126				
121	AAGGAGATCGTCGGCGCTACAACTGGGTGAGTACATCGGACATCTACGACAAAGTG	180				
127	AAGGAGATCGTCGGCGCTACAACTGGGTGAGTACATCGGACATCTACGACAAAGTG	186				
181	TCGGGGGACATCGACAAAGAGGTCGACGTGAGTGTCTGGGGGGGGGGCCATCTCC	240				
187	TCGGGGGACATCGACAAAGAGGTCGACGTGAGTGTCTGGGGGGGGGGCCATCTCC	246				
241	CACCAAGTCGAGACAAAGAAATTCAGGTATGCGGTATTCATGGCTATGTCTGGCC	300				
247	CACCAAGTCGAGACAAAGAAATTCAGGTATGCGGTATTCATGGCTATGTCTGGCC	306				
301	CAGCAGCCATTTCACTGAGAAAAATCAAGCCAAAGTACCCGAGCTACGAGCTACCTGG	360				
307	CAGCAGCCATTTCACTGAGAAAAATCAAGCCAAAGTACCCGAGCTACGAGCTACCTGG	366				
361	GCTAAGCAGGCTAC 375					
367	GCTAAGCAGGCTAC 381					
RESULT 5						
BMA71874						

LOCUS BM471874 578 bp mRNA linear EST 05-FEB-2002
 DEFINITION AGENCOURT_6464398 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5581585
 5', mRNA sequence.
 ACCESSION BM471874
 VERSION BM471874.1 GI:18520916
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 578)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: scapbs@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone Distribution: MGC clone distribution information
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
 Plate: LLM12341 row: 1 column: 02
 High quality sequence stop: 577.
 Location/Qualifiers
 1..578
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5581585"
 /clone_1ib="NIH-MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally. Primer: oligo dr.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 BASE COUNT 128 a 179 c 166 g 105 t
 ORIGIN

Query Match 100.0%; Score 375; DB 13; Length 578;
 Best Local Similarity 100.0%; Pred. No. 5,76-86;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTGGGGGACCTGCTCTATCTGATGGGACATCGACCTCGAGCGGCTTC 60
 DB 12 ATGGCGTGGGGGACCTGCTCTATCTGATGGGACATCGACCTCGAGCGGCTTC 60
 QY 61 AAGTATGCTGATCGGAGTCGACCTCGGCTCCGCGGCTCGGCTCGAGAGAGC 120
 DB 72 AAGTATGCTGATCGGAGTCGACCTCGGCTCCGCGGCTCGGCTCGAGAGAGC 120
 QY 121 AAGGAGATGCTGGGGGCTACAGTGGGCTGATACCATCGGACATCTACGACAAAGT 180
 DB 132 AAGGAGATGCTGGGGGCTACAGTGGGCTGATACCATCGGACATCTACGACAAAGT 180
 QY 181 TCGGGGACATCAGAACAGAGGCTGAGTGTGGGCGGCGGCGCATCTCC 240
 DB 192 TCGGGGACATCAGAACAGAGGCTGAGTGTGGGCGGCGGCGCATCTCC 240
 QY 241 CACGAGATCGAGACAGAGATTCACGTAGCGCTATTCATGGCCCTATGCTCGCC 300
 DB 252 CACGAGATCGAGACAGAGATTCACGTAGCGCTATTCATGGCCCTATGCTCGCC 300
 QY 301 CAGCAGCGCATTTCAATGAGAAATCAAGCAAGTACCCGAGTACGAGTACCTGG 360
 DB 312 CAGCAGCGCATTTCAATGAGAAATCAAGCAAGTACCCGAGTACGAGTACCTGG 360
 QY 361 GCTAAGAGCGGCTAC 375
 DB 372 GCTAAGAGCGGCTAC 386

RESULT 6
 LOCUS BM691037
 DEFINITION BM691037 597 bp mRNA linear EST 28-FEB-2002
 UI-E-C11-aba-h-02-0-UI-1 UI-E-C11 Homo sapiens cDNA clone
 ACCESSION UI-E-C11-aba-h-02-0-UI 5', mRNA sequence.
 VERSION BM691037
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 597)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Nat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 577-597, >AT rich flow-complexity
 Seq primer: M13 Reverse
 Location/Qualifiers
 1..597
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-aba-h-02-0-UI"
 /clone_1ib="UI-E-C11"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site: 1: EcoR I; Site: 2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA,
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into p773-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 (dt)18 tail. The sequence tag for this library is ACCCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."
 BASE COUNT 135 a 186 c 171 g 105 t
 ORIGIN

Query Match 100.0%; Score 375; DB 14; Length 597;
 Best Local Similarity 100.0%; Pred. No. 5,86-86;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTGGGGGACCTGCTCTATCTGATGGGACATCGACCTCGAGCGGCTTC 60
 DB 22 ATGGCGTGGGGGACCTGCTCTATCTGATGGGACATCGACCTCGAGCGGCTTC 60
 QY 61 AAGTATGCTGATCGGAGTCGACCTCGGCTCCGCGGCTCGGCTCGAGAGAGC 120
 DB 82 AAGTATGCTGATCGGAGTCGACCTCGGCTCCGCGGCTCGGCTCGAGAGAGC 141

`/dev-stgase="adult" (T1 phage resistant) /
/lab-host=BDH10 (Life Technologies) /
/note=Organ: eye; Vector: p1773-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
I-E-DWG is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Ronald, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not`

QY	1	ATGGGCGTGCGGACCTCGCTTCATTCCTGATGTGGACATCGACTCCAGCGGCGCTTC	80
Db	15	ATGGCGGCGGGGAGACCTCGCTTCATTCCTGATGTGGACATCGACTCCAGCGGCGCTTC	74
QY	61	AAGTATGTGCGTATCCGAGCTCCACTCGGCTCCCGGCTCCGGGCTCCGGCTCGAGAGAC	120
Db	75	AAGTATGTGCTATCCGAGCTCCACTCCGCTCCCGGCTCCGGGCTCCGGCTCGAGAGAC	134
QY	121	AAGGAGATCGTCGCGGGCTACAACTGGGCTAAGTACCATGGAGCATCTACGACAAAGTG	180
Db	135	AAGGAGATCGTCGCGGGCTACAACTGGGCTAAGTACCATGGAGCATCTACGACAAAGTG	194
QY	181	TCGGCGCATGTCAGAAAGAAAGGCGCGACGTGAGTGTCTGGCGGGCGGGCGCATCTCC	240
Db	195	TCGGCGCATGTCAGAAAGAAAGGCGCGACGTGAGTGTCTGGCGGGCGGGCGCATCTCC	254
QY	241	CACGACGATCAGACAAAGAAAGATTACGTGTCGCTATTCATGAGCCTATAGTGCCTGGCC	300
Db	255	CACGACGATCAGACAAAGAAAGATTACGTGTCGCTATTCATGAGCCTATAGTGCCTGGCC	314
QY	301	CAGCAGCCATTCCACTGAGAAATCAAAAGCAAGTACCCGACTACGAGTCACTGG	360
Db	315	CAGCAGCCCAATTCCTCACTGAGAAATCAAAAGCAAGTACCCGACTACGAGTCACTGG	374
QY	361	GCTAACGACGGCTAC	375
Db	375	GCTAACGACGGCTAC	389

RESULT 8	BI833178	604 bp	mrna	linear	EST 04-OCT-2001
LOCUS	60308721121 NIH_MGC_120		Homo sapiens	CDNA clone	IMAGE:5226219 5'
DEFINITION	BI833178		mrna sequence.		
ACCESSION	BI833178				
VERSION	BI833178.1	GI:15944728			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/				
AUTHORS	1 (bases 1 to 604)				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				
	robertv@mail.nih.gov				

FEATURES	Location/Qualifiers
source	1. .604
	4. .604
	5. .604
	6. .604
	7. .604
	8. .604
	9. .604
	10. .604
	11. .604
	12. .604
	13. .604
	14. .604
	15. .604
	16. .604
	17. .604
	18. .604
	19. .604
	20. .604
	21. .604
	22. .604
	23. .604
	24. .604
	25. .604
	26. .604
	27. .604
	28. .604
	29. .604
	30. .604
	31. .604
	32. .604
	33. .604
	34. .604
	35. .604
	36. .604
	37. .604
	38. .604
	39. .604
	40. .604
	41. .604
	42. .604
	43. .604
	44. .604
	45. .604
	46. .604
	47. .604
	48. .604
	49. .604
	50. .604
	51. .604
	52. .604
	53. .604
	54. .604
	55. .604
	56. .604
	57. .604
	58. .604
	59. .604
	60. .604
	61. .604
	62. .604
	63. .604
	64. .604
	65. .604
	66. .604
	67. .604
	68. .604
	69. .604
	70. .604
	71. .604
	72. .604
	73. .604
	74. .604
	75. .604
	76. .604
	77. .604
	78. .604
	79. .604
	80. .604
	81. .604
	82. .604
	83. .604
	84. .604
	85. .604
	86. .604
	87. .604
	88. .604
	89. .604
	90. .604
	91. .604
	92. .604
	93. .604
	94. .604
	95. .604
	96. .604
	97. .604
	98. .604
	99. .604
	100. .604

[illegible]

QY	1	ATGGCGTGGGGACCTCGCTCTCTATTCTCGATGTGGACATCGATCCGACGGCGCTTTC	60
Db	15	ATGGCGCTGGCGACCTCGCTCTCTATTCTCGATGTGGACATCGATCCGACGGCGCTTTC	74
QY	61	AAGTATGTGCTGATGCCAGTCCACTCGGGCTCCCGGCTCCGGGGCTCCGGGCTGCAGAGAGC	120
Db	75	AAGTATGTGCTGATGCCAGTCCACTCGGGCTCCCGGCTCCGGGGCTCCGGGCTGCAGAGAGC	134
QY	121	AAGGAGATCGTGGCGCGCTCAAGTGGGGGTGATGCCATTCGGGCACTATCGACAAAGTG	180
Db	135	AAGGAGATCGTGGCGCGCTCAAGTGGGGGTGATGCCATTCGGGCACTATCGACAAAGTG	194
QY	181	TGGGGCGCATGCGAAGCAAGGCTGCGACTGTGATGTCTGGGGCGGGCGGCACATTCGC	240
Db	195	TGGGGCGCATGCGAAGCAAGGCTGCGACTGTGATGTCTGGGGCGGGCGGCACATTCGC	254
QY	241	CACGAGGTGAGGACAAAGATTACGTGTACGGCTATTCCATATGGGCTATATGGTCTCTGCC	300
Db	255	CACGAGGTGAGGACAAAGATTACGTGTACGGCTATTCCATATGGGCTATATGGTCTCTGCC	314
QY	301	CAGCAGCCATTTCACACTGAGAAATCAAGACCCAGATACCGGACTACGAGGTCACTGG	360
Db	315	CAGCAGCCATTTCACACTGAGAAATCAAGACCCAGATACCGGACTACGAGGTCACTGG	374
QY	361	GCTACGAGGCGCTAC	375
Db	375	GCTACGAGGCGCTAC	389

RESULT 9	
BG819649	
LOCUS	BG819649
DEFINITION	607 bp mRNA linear EST 22-MAY-2001 602783332FI NCI_CGAP_Brn67 Homo sapiens CDNA clone IMAGE:4934238
ACCESSION	BG819649
	5', mRNA sequence.

KEYWORDS EST.
SOURCE human.

ORGANISM

REFERENCE 1 (bases 1 to 607) Baccharis; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
TITLE
NIH-MGC <http://mgc.>

THE
National Institutes of Health,
JOURNAL
Unpublished (1988)

COMMENT

Revised (1999)

Contact: Robert Strausberg Ph.D.

Email: cgapbs-r@mail.nih.gov

Issue Procurement: David N. Louis, M.D.
CDNA Library Procurement: [redacted]

Library Preparation: Lite Technologies, Inc
CDNA Library Arraved by: The I M a C T

DNA Sequencing by: Incyte Genomics, Inc.
Consortium (LNL),

found through the TWC

... the I.M.A.G.E. Consortium/LINTL at

```

http://image.lnl.gov
plate: L1AM10862 row: 0 column: 07
High quality sequence stop: 576.
location/Qualifiers
1. 607
FEATURES
source
```

BASE COUNT
ORIGIN

134 a 195 c 172 g 106 t

"/seq=ccat-unibw (11 phage-reslant)"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dr.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

	Query Match	Best Local Similarity	100.0%;	Score 375;	DB 12;	Length 607;
	Matches 375;	Conservative	100.0%;	Pred. No. 5,8e-86;	Mismatches 0;	Indels 0;
QY	1	ATAGCGGCGTGGCGGACCTCGCTCTCATTTCCGATGTGAGCATTCGACCTCCGAGGGCGTTC	60			
Db	22	ATAGGCGGTGGGCGACCTCGCTCTCATTTCCGATGTGAGCATTCGACCTCCGAGGGCGTTC	60			
QY	61	AAGTATGTGCGTATCCGATCCGATCGCTGCCCTCCCGGGGCTCCGGGCTCGAGAGAC	120			
Db	82	AAGTATGTGCGTATCCGATCCGATCGCTGCCCTCCCGGGGCTCCGGGCTCGAGAGAC	141			
QY	121	AAGGAGATCGTGGCGGCTACAAAGTGGGTGTGATACCATGCGGCATCTACGACAAAGTG	180			
Db	142	AAGGAGATCGTGGCGGCTACAAAGTGGGTGTGATACCATGCGGCATCTACGACAAAGTG	201			
QY	181	TGGGGGCACATCGAGAAAGCGCTGCGACCTGATGTCTGGGCGGGCGGGCATCTCC	240			
Db	202	TGGGGGCACATCGAGAAAGCGCTGCGACCTGATGTCTGGGCGGGCGGGCATCTCC	261			
QY	241	CACGAGAGTCAGACAAAGAAATTCAGCTGTAGCGCTATTCCATGTGCCCTATGTCTCTGCC	300			
Db	262	CACGAGAGTCAGACAAAGAAATTCAGCTGTAGCGCTATTCCATGTGCCCTATGTCTCTGCC	321			
QY	301	CACACAGCCATTTCAACTGAGAAAATCAAAGCCAAAGACCCGAGCTACGAGGTCACTGG	360			
Db	322	CACACAGCCATTTCAACTGAGAAAATCAAAGCCAAAGTACCCGACTACGAGGTCACTGG	381			
QY	361	GCTAACGAGAGGCTTC	375			
Db	382	GCTAACGAGAGGCTTC	396			

RESULT 10
BM918530

LOCUS
DEFINITION

AGENCOURT_6708531

Accession BM918530

VERSION	BM918530.1	GI:193
KEYWORDS	EST	

SOURCE
human.

ORGANISM

Eukaryota; Metazoa
Mammalia: Mammalia

REFERENCE
1 (bases 1 to 612)

AUTHORS NIH-MGC <http://mgc>

THE
National Institute
JOURNAL
Published 1999

COMMENT
Contact: Robert St

Email: cyapbs-r@ma

Issue Procurement
CDNA Library Proc.

CDNA Library Array

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DTIP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM2431 row: m column: 07
High quality sequence stop: 510.

FEATURES
source
location/Qualifiers
1..1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6265134"
/clone_lib="NIH-MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(C). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 238 a 312 c 301 g 155 t
ORIGIN

Query Match 100.0%; Score 375; DB 14; Length 1006;
Best Local Similarity 100.0%; Pred. No. 6.9e-86;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGCGGACCTCGCTCATCTCTGATGTGACATGACATCCGACGGCTTC 60
Db 8 ATGGCGGTGGCGGACCTCGCTCATCTCTGATGTGACATGACATCCGACGGCTTC 67
QY 61 AAGTATGTGCTGATCCGAGTCACACTCGGCTCCCGCTCCGGGGCTCCGGCTTC 120
Db 68 AAGTATGTGCTGATCCGAGTCACACTCGGCTCCCGCTCCGGGGCTCCGGCTTC 127
QY 121 AAGGAGATGCTGGCGGCTACAGTGGGCTAGTACATGCGGACATCTACGACAAAGT 180
Db 128 AAGGAGATGCTGGCGGCTACAGTGGGCTAGTACATGCGGACATCTACGACAAAGT 187
QY 181 TCGGGGACATGACAGAAAGGCTGACATGCTGAGTGTGGGGCGGGCGCATCTCC 240
Db 188 TCGGGGACATGACAGAAAGGCTGACATGCTGAGTGTGGGGCGGGCGCATCTCC 247
QY 241 CACGAGATCAGAGACAGAAAGTACGTGACGCTATTCATGCGCTATGCTTCGCC 300
Db 248 CACGAGATCAGAGACAGAAAGTACGTGACGCTATTCATGCGCTATGCTTCGCC 307
QY 301 CAGCAGGCAATTCACATGAGAAATCAAAAGCAAGTACCCGACATACGAGTCACTGG 360
Db 308 CAGCAGGCAATTCACATGAGAAATCAAAAGCAAGTACCCGACATACGAGTCACTGG 367
QY 361 GCTAACGACGGCTAC 375
Db 368 GCTAACGACGGCTAC 382

RESULT 13
BF345224

LOCUS BF345224 1010 bp mRNA linear EST 22-NOV-2000
DEFINITION 602018107F1 NCL-CGAP_Brn67 Homo sapiens CDNA clone IMAGE:4153865
5', mRNA sequence.
ACCESSION BF345224
VERSION BF345224.1 GI:11292725
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LAM9422 row: k column: 18
High quality sequence stop: 593.

FEATURES
source
location/Qualifiers
1..1010
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4153865"
/clone_lib="NCL-CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL-CGAP Library."

BASE COUNT 312 a 289 c 282 g 127 t
ORIGIN

Query Match 100.0%; Score 375; DB 12; Length 1010;
Best Local Similarity 100.0%; Pred. No. 6.9e-86;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGGCGGACCTCGCTCATCTCTGATGTGACATGACATCCGACGGCTTC 60
Db 5 ATGGCGGTGGCGGACCTCGCTCATCTCTGATGTGACATGACATCCGACGGCTTC 64
QY 61 AAGTATGTGCTGATCCGAGTCACACTCGGCTCCCGCTCCGGGGCTCCGGCTTC 120
Db 65 AAGTATGTGCTGATCCGAGTCACACTCGGCTCCCGCTCCGGGGCTCCGGCTTC 124
QY 121 AAGGAGATGCTGGCGGCTACAGTGGGCTAGTACATGCGGACATCTACGACAAAGT 180
Db 125 AAGGAGATGCTGGCGGCTACAGTGGGCTAGTACATGCGGACATCTACGACAAAGT 184
QY 181 TCGGGGACATGACAGAAAGGCTGACATGCTGAGTGTGGGGCGGGCGCATCTCC 240
Db 185 TCGGGGACATGACAGAAAGGCTGACATGCTGAGTGTGGGGCGGGCGCATCTCC 244
QY 241 CACGAGATCAGAGACAGAAAGTACGTGACGCTATTCATGCGCTATGCTTCGCC 300
Db 245 CACGAGATCAGAGACAGAAAGTACGTGACGCTATTCATGCGCTATGCTTCGCC 304
QY 301 CAGCAGGCAATTCACATGAGAAATCAAAAGCAAGTACCCGACATACGAGTCACTGG 360
Db 305 CAGCAGGCAATTCACATGAGAAATCAAAAGCAAGTACCCGACATACGAGTCACTGG 364
QY 361 GCTAACGACGGCTAC 375
Db 365 GCTAACGACGGCTAC 379

RESULT 14
BF405028/c 579 bp mRNA linear EST 28-NOV-2000
LOCUS BF405028
DEFINITION BF405028
ACCESSION BF405028
VERSION BF405028.1 GI:11393003
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 579)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA library preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-Yes.
Location/Qualifiers
1..579
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UR-CAL-bio-g-01-0-UI"
/clone_1lb="UR-CAL"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The UR-CAL
library is a subtracted library derived from the following
tissues: thalamus, cerebellum, hypothalamus, medulla, pons
, midbrain, cerebral cortex, corpus striatum, testis, and
hippocampus. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratel.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_SEO=None found"
BASE COUNT 105 a 165 c 183 g 126 t
ORIGIN
Query Match 99.6%; Score 373.4; DB 12; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.5e-85;
Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 392 TCGGGGACATGACAGAAAGGCTGCGACTGTGATGTCTGGGGGGGCGCATCTCC 333
OY 241 CACCAGAGTCAGACAAAGAAATTCACGCTACGGCTATTCATGATGCTCC 300
Db 332 CACCAGAGTCAGACAAAGAAATTCACGCTACGGCTATTCATGATGCTCC 273
OY 301 CAGCAGCCATTCACACTGAGAAATCAAGCCAGTACCCGACTACGAGTCTCG 360
Db 272 CAGCAGCCATTCACACTGAGAAATCAAGCCAGTACCCGACTACGAGTCTCG 213
OY 361 GCTAACGACGCTAC 375
Db 212 GCTAACGACGCTAC 198

RESULT 15
B0651896 589 bp mRNA linear EST 15-JUL-2002
LOCUS B0651896
DEFINITION AGENCOURT_8355810 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6285251
5', mRNA sequence.
ACCESSION B0651896
VERSION B0651896.1 GI:21776068
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 589)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1995)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@imail.nih.gov
Tissue Procurement: CGAP (Stanford)
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10M2484 row: C column: 12
High quality sequence start: 2
High quality sequence stop: 588.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6285251"
/clone_1lb="NIH_MGC_100"
/tissue.type="hepatocellular carcinoma, cell line"
/note="organ: liver; Vector: pOT81; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(g). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the Laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 136 a 182 c 164 g 107 t
ORIGIN
Query Match 99.6%; Score 373.4; DB 14; Length 589;
Best Local Similarity 99.7%; Pred. No. 1.5e-85;
Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGGCGGTGGGCGACTGCTCTCATCTTCGATGCGACATGCGAGCGGCTTC 60
Db 8 ATGGCGGTGGGCGACTGCTCTCATCTTCGATGCGACATGCGAGCGGCTTC 67
OY 61 AAGTATGTGATCGAGTCCACTGCGCTCCCGGCTCCGGGCTCCAGAGAGC 120

Db 68 |||||
AAGTATGCTGATCCGAGTCCACTGGCTCCCGGCTCCGGGGCTCCGGGCTCGAGAGGC 127
QY 121 AAGGATCTGTCGGCGGCTACAGTGGGCTGAGTACCATCGGACATCTACGACAAAGTG 180
Db 128 AAGGATCTGTCGGCGGCTACAGTGGGCTGAGTACCATCGGACATCTACGACAAAGTG 187
QY 181 TCGGGCCGACATGACAAAGCAAGGCTGCACTGTGAGTGTCTGGGGCGGGGGCGCATCTGC 240
Db 188 TCGGGCCGACATGACAAAGCAAGGCTGCACTGTGAGTGTCTGGGGCGGGGGCGCATCTGC 247
QY 241 CACGAGATGAGGACAGAGATTCAGTGTACGGCTATTCCATGAGCTATGCTCTGCC 300
Db 248 CACGAGATGAGGACAGAGATTCAGTGTACGGCTATTCCATGAGCTATGCTCTGCC 307
QY 301 CAGCAGCCATTTCACCTGAGAAAATCAAGCCAGTACCCGACTAGAGAGTCACTGC 360
Db 308 CAGCAGCCATTTCACCTGAGAAAATCAAGCCAGTACCCGACTAGAGAGTCACTGC 367
QY 361 GCTAAGCAGCGCTAC 375
Db 368 GCTAAGCAGCGCTAC 382

Search completed: May 13, 2003, 19:23:05
Job time : 1760 secs

102 GGCTCCGGCTCAAGACGCAAGATCGTCGGCGCTACAAAGTGGAGCTATACCAATC 161

Db 363 CGAGCCGCTGTGTGACATGACGACATGACATGACGACCAATGCGCCCTACCACTT 422
QY 162 GCACATCTACGACAAAGTGTGCGGACATGACGACAAAGCAAGCTGCGCA 209
Db 423 CGACCTGCGCGGACGCGCTGTGCGGACATGCGGACGACAAAGGCGGAGGA 470

RESULT 4

US-09-007-005-17
Sequence 17, Application US/0907005B

Patent No. 6258558

GENERAL INFORMATION:

APPLICANT: Szostak, Jack W.

APPLICANT: Roberts, Richard W.

APPLICANT: Liu, Rihue

TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN

TITLE OF INVENTION: FUSIONS

FILE REFERENCE: 00786/350003

CURRENT APPLICATION NUMBER: US/09/007,005B

CURRENT FILING DATE: 1998-01-14

EARLIER APPLICATION NUMBER: 60/035,963

EARLIER FILING DATE: 1997-01-27

EARLIER APPLICATION NUMBER: 60/064,491

EARLIER FILING DATE: 1997-11-06

NUMBER OF SEQ ID NOS: 33

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17

LENGTH: 289

TYPE: RNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Translation template

NAME/KEY: misc-feature

LOCATION: (1)...(289)

OTHER INFORMATION: n = A,T,C or G

US-09-007-005-17

Query Match 9.1%; Score 34.2; DB 4; Length 289;

Best Local Similarity 2.9%; Pred. No. 1;

Matches 5; Conservative 82; Mismatches 88; Indels 0; Gaps 0;

QY 116 AGAGCAAGAGATCGTGGCGGCTACAGAGTGGCTGAGTACATGCGGACATGACGACA 175

Db 55 RGRNR 114

QY 176 AAGTGTGCGGACATGACGACAAAGCAAGCTGCGACTGTGAGTGTCTGGCGCGGCGCA 235

Db 115 RSRNR 174

QY 236 TCTCCACGACAGATGACGACAAAGATTCAGCTGACGCTATTCCATGCGCTA 290

Db 175 RSRNR 229

RESULT 5

US-09-244-796-17

Sequence 17, Application US/09244796

Patent No. 6281344

GENERAL INFORMATION:

APPLICANT: Szostak, Jack W.

APPLICANT: Roberts, Richard W.

APPLICANT: Liu, Rihue

TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN

TITLE OF INVENTION: FUSIONS

FILE REFERENCE: 00786/350007

CURRENT APPLICATION NUMBER: US/09/244,796

CURRENT FILING DATE: 1999-02-05

EARLIER APPLICATION NUMBER: 60/035,963

EARLIER FILING DATE: 1997-01-27

EARLIER APPLICATION NUMBER: 60/064,491

EARLIER FILING DATE: 1997-11-06

EARLIER APPLICATION NUMBER: 09/007,005

EARLIER FILING DATE: 1998-01-14

NUMBER OF SEQ ID NOS: 33

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17

LENGTH: 289

TYPE: RNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Translation template

NAME/KEY: misc-feature

LOCATION: (1)...(289)

OTHER INFORMATION: n = A,T,C or G

US-09-244-796-17

Query Match 9.1%; Score 34.2; DB 4; Length 289;

Best Local Similarity 2.9%; Pred. No. 1;

Matches 5; Conservative 82; Mismatches 88; Indels 0; Gaps 0;

QY 116 AGAGCAAGAGATCGTGGCGGCTACAGAGTGGCTGAGTACATGCGGACATGACGACA 175

Db 55 RGRNR 114

QY 176 AAGTGTGCGGACATGACGACAAAGCAAGCTGCGACTGTGAGTGTCTGGCGCGGCGCA 235

Db 115 RSRNR 174

QY 236 TCTCCACGACAGATGACGACAAAGATTCAGCTGACGCTATTCCATGCGCTA 290

Db 175 RSRNR 229

RESULT 6

US-08-402-282-3

Sequence 3, Application US/08402282

Patent No. 5476768

GENERAL INFORMATION:

APPLICANT: Pearson, Robert E.

APPLICANT: Dickson, Julie A.

APPLICANT: Hamilton, Paul T.

APPLICANT: Little, Michael C.

APPLICANT: Beyer, Jr., Wayne F.

TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE

TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS: Richard J. Rodrick, Becton Dickinson and

ADDRESSEE: Company

STREET: 1 Becton Drive

CITY: Franklin Lakes

STATE: NJ

COUNTRY: US

ZIP: 07417

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/402,282

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fugit, Donna R.

REGISTRATION NUMBER: 32,135

REFERENCE/DOCKET NUMBER: P-3283

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 15664 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

[illegible]

APPLICANT: Pearson, Robert E.
APPLICANT: Dickson, Julie A.
APPLICANT: Hamilton, Paul T.
APPLICANT: Little, Michael C.
APPLICANT: Beyer Jr., Wayne F.
TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
ADDRESSEE: Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/508,004
APPLICATION NUMBER: US/08/508,004
FILING DATE: 27-JUL-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/402,282
FILING DATE: 10-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3283
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 15664 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 222..425
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 451..747
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 747..1109
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1109..2014
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2034..2747
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2747..3109
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3109..3444
OTHER INFORMATION: /function="potential open reading

OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3444..3728
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3731..4855
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4855..5376
OTHER INFORMATION: /function="potential coding
OTHER INFORMATION: sequence"
OTHER INFORMATION: /product="L5 gp37 homolog"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5382..5747
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5837..6307
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6403..7770
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7770..8006
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8033..8236
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8244..9443
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9450..10244
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10371..10586
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11115..11786
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function="potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14499..15664
OTHER INFORMATION: /function="potential open reading

NAME/KEY: misc.feature
LOCATION: 1471..15154
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 15154..15426
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 15429..15664
OTHER INFORMATION: /function="potential open reading"
OTHER INFORMATION: frame"
US-08-508-004-3

Query Match 9.0%; Score 33.8; DB 1; Length 15664;
Best Local Similarity 52.5%; Pred No. 3.5;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

OY 149 CTGAGTACCATGCGGACATCTACGACAAAGTGTGGGCGCATGCGAGAGCAAGGCTGCG 208
DB 7151 CCGAGAGCTGGGGGCGGAGGTGGCGGCAAGCGTCTCGACACGCTGCACAAAGGCGCG 7210
OY 209 ACTGTAGTGTCTGGGCGCGCGCATCTCCACAGATCAGACAGCAAGATTCACG 268
DB 7211 AGTTCAGCGCGGCGCGGCGGTGGGCGCACCCCTCAAGCGCGCTTCGAGCGATCGTGG 7270
OY 269 TGTAGCGCTATTCATGCGCCT 289
DB 7271 TGGCCGCAACCGATGCGCT 7291

RESULT 8
US-08-402-066-3
Sequence 3, Application US/08402066
Patent No. 5612182
GENERAL INFORMATION:

APPLICANT: Pearson, Robert E.
APPLICANT: Dickson, Julie A.
APPLICANT: Hamilton, Paul T.
APPLICANT: Little, Michael C.
APPLICANT: Bayer Jr., Wayne F.
TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
NUMBER OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
CORRESPONDENCE ADDRESSES: 6

ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
ADDRESS: Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/402,066

FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3283
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 15664 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 222..425
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 451..747
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 747..1109
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1109..2014
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 2034..2747
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 2747..3109
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3109..3444
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3444..3728
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3731..4855
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 4855..5376
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 5382..5747
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 5837..6307
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 6403..7770
OTHER INFORMATION: /function="potential open reading"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 7770..8006
OTHER INFORMATION: /function="potential open reading"
FEATURE:

NAME/KEY: misc_feature
LOCATION: 8033..8236 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8244..9443 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9450..10244 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10371..10586 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11115..11786 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14771..15154 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15154..15426 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15429..15664 /function="potential open reading"
OTHER INFORMATION: frame"
US-08-402-066-3
Query Match 9.0%; Score 33.8; DB 1; Length 15664;
Best Local Similarity 52.5%; Pred. No. 3.5; Mismatches 67; Indels 0; Gaps 0;
Matches 74; Conservative 0;
QY 149 CTGAGTACCATCTAGACATCTAGCAAGTGTGGGGCATGAGCAAGAGCTTGGC 208
DB 7151 CCGAGGACTGGGGGCAAGGTGGGCAAAACGGTCTCTGACACGCTGACAGGGCGGCG 7210
QY 209 ACTGTGAGTCTGGGGGGGGGGCATCTCCACACAGAGTCAGAGCAGAGATTCAAG 268
DB 7211 AGTTGCGAGCGGGCGGCGCTGGGGCACACCTCAAGCGGGCGTTCCAGCCGATCGTGG 7270
QY 269 TGTACGGCTATTCATGGCCT 289
DB 7271 TGGCCCGCAACGATGCCGT 7291

RESULT 9
US-08-402-068-3
Sequence 3, Application US/08402068
Patent No. 5633159
GENERAL INFORMATION:

APPLICANT: Pearson, Robert E.
APPLICANT: Dickson, Julie A.
APPLICANT: Hamilton, Paul T.
APPLICANT: Little, Michael C.
APPLICANT: Beyer Jr., Wayne F.
TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
ADDRESS: Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/402,068
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3283
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 15664 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 222..425 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 451..747 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 747..1109 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1109..2014 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2034..2747 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2747..3109 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3109..3444 /function="potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature

LOCATION: 3444..3728
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 3731..4855
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 4855..5376
OTHER INFORMATION: /function= "potential coding
OTHER INFORMATION: sequence"
FEATURE: /product= "L5 gp37 homolog"
NAME/KEY: misc_feature
LOCATION: 5382..5747
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 5837..6307
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 6403..7770
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 7770..8006
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 8033..8236
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 8244..9443
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 9450..10244
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 10371..10586
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 11115..11786
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 14771..15154
OTHER INFORMATION: /function= "potential open reading

OTHER INFORMATION: frame"
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 15154..15426
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
NAME/KEY: misc_feature
LOCATION: 15429..15664
OTHER INFORMATION: /function= "potential open reading
FEATURE: frame"
US-08-402-068-3
Query Match 9.0%; Score 33.8; DB 1; Length 15664;
Best Local Similarity 52.5%; Pred. No. 3.5;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 149 CTGAGTACCATGCGGACATCTACGACAAAGTGTGCGGAGCATGCGACGACGAGTGGC 208
Db 7151 CCGAGGACTGGGGCGGACGAGGTGGCGCAACGGTCTCTGACACGCTGACGAGGCGCGC 7210
QY 209 ACTGTAGTGTGTGGGGCGGCGCATCTCCACACAGTACAGACGACGACGATTCACG 268
Db 7211 AGTTGAGCGCGCGCGCGGTGGGCGACACCTCAAGCGCGGCTTGAGCGCATCTGG 7270
QY 269 TGTAGCGCTATTCCATGGCCT 289
Db 7271 TGGCCCGCAACCGATGCCGT 7291

RESULT 10
US-08-804-227C-7
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhloss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYMERIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4437 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002

FEATURE: CDS
NAME/KEY: 14046..20036
LOCATION: 14046..20036
FEATURE: CDS
NAME/KEY: 20110..31284
LOCATION: 20110..31284
FEATURE: CDS
NAME/KEY: 31329..36071
LOCATION: 31329..36071
FEATURE: CDS
NAME/KEY: 36155..41830
LOCATION: 36155..41830
US-08-804-227C-7

Query Match 8.9%; Score 33.4; DB 2; Length 44377;
Best Local Similarity 46.1%; Pred. No. 5.8;
Matches 112; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 59 TCAGTATGCTGATCGAGTCCAGTCCGCTCCCGCTCCGGGCTCCGGCTGCAGAGA 118
DB 19342 TCCTGCGTACCTGCGCGGCTACCGCGGAGCGCCGACACGCGCGGGCGCGGCGG 19401

QY 119 GCAAGGATGCTGCGCGGCTACAGTGGGCTGAGTACATCGGACATCTACGACAAAG 178
DB 19402 GCACCGGGGAGAGAGAGCGTCCGCTGCGCTCCCGCGCGCGCGCGCGGCGGGA 19461

QY 179 TGTGCGGCGACATGCAAGAGCGCTGCGACTGTGAGTGTCTGGGCGCGGCGCATCT 238
DB 19462 CGCTGCGCGCGCGGCTGCGCGGAGCTTCGCGACCGGACGCGCTCTCTGCTGACG 19521

QY 239 CCCACGAGTACGACAAAGATTCAGTGTACGCTATTCATGCGCTATGCTCTG 298
DB 19522 TGTGCGGCGAGCGAGTCTGCGGCGTGTGCGACACGCGCGCGCGCGCATCTGCGCG 19581

QY 299 CCC 301
DB 19582 CCC 19584

RESULT 11
US-08-804-198-1
Sequence 1, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-198-1

Query Match 8.9%; Score 33.4; DB 2; Length 44377;
Best Local Similarity 46.1%; Pred. No. 5.8;
Matches 112; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 59 TCAGTATGCTGATCGAGTCCAGTCCGCTCCCGCTCCGGGCTCCGGCTGCAGAGA 118
DB 19342 TCCTGCGTACCTGCGCGGCTACCGCGGAGCGCCGACAGCGCGCGGGCGCGGCGG 19401

QY 119 GCAAGGATGCTGCGCGGCTACAGTGGGCTGAGTACATTCGCGACATCTACGACAAAG 178
DB 19402 GCACCGGGGAGAGAGCGTCCGCTGCGCTCCCGCGCGCGCGCGCGCGGCGGGA 19461

QY 179 TGTGCGGCGACATGCAAGAGCGCTGCGACTGTGAGTGTCTGGGCGCGGCGCATCT 238
DB 19462 CGCTGCGCGCGCGGCTGCGCGGAGCTTCGCGACCGGACGCGCTCTCTGCTGACG 19521

QY 239 CCCACGAGTACGACAAAGATTCAGTGTACGCTATTCATGCGCTATGCTCTG 298
DB 19522 TGTGCGGCGAGCGAGTCTGCGGCGTGTGCGACACGCGCGCGCGCGCATCTGCGCG 19581

QY 299 CCC 301
DB 19582 CCC 19584

RESULT 12
US-07-887-072B-1
Sequence 1, Application US/07887072B
Patent No. 5424191
GENERAL INFORMATION:
APPLICANT: Prasad Ph.D., Gaddamanugu L.
APPLICANT: Cooper M.D., Herbert L.
TITLE OF INVENTION: EPITHELIAL CELL SPECIFIC DIFFERENTIATION
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 W. Broadway, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/887,072B

FILED DATE: 20-MAY-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kirkpatrick Ph.D., Anita M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH021.021A
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-887-072B-1

Query Match
Best Local Similarity 8.7%; Score 32.8; DB 1; Length 1245;
Matches 161; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

QY 2 TGGGGTGGGGGACCTGCTCTATTCCTGATGTGACATTCGACGCGCTTTCA 61
DB 309 TGGGGCTGCTGGACACCCATCTATCAAGAGGCGCGGAGCCGAGAGCCGGGTCTTC 368
QY 62 AGTATGCTGATCCAGTCCACTGCGTCCCGCTCCGGGCTCCGGGTGACAGAGA 121
DB 369 ACCTGAAGATGAAGGGTGTACTTACCTGCTGCTGCGGAGGTGACACCGGTGACGACA 428
QY 122 AGAGATCGTGGGGCTTCAAGTGGGCTGATACCATGCGGACATTCAGCAAAAGTGT 181
DB 429 AGAAGGCGATCATTTGACTAGCCCGGTGACCTACCAAGAGGCGCATGTGACATCAGCAAGA 488
QY 182 CGGGGACATGCAAGAGCAAGGCTGCACTGTGATGTGTGGGGGGGGGGGCAATCC 241
DB 489 AG---GAGATGCCGCCACCAACCCCACTCCGCTGGGCTGCGCTGAACTTTTCCGT 545
QY 242 ACCAGAGTGAAGCAAGAAATTCACGTGTACGCTATTCATGAGGCTTATGCTGCC 301
DB 546 TCCACTAGAGATGCGCAACAGCCCGGAGAGGCGCATCTCTGCGCAAGACCACTTTCG 605
QY 302 AGCAGCCATTTCACCTGAGAAATCAAGCCAGTACCCGCGCTAGAGAGTCC 357
DB 606 ACGAGGCGATGCTGATCTGCAACCCCTCAGCGAGGACTCTTACAAAGACAGACCC 661

RESULT 13
US-08-466-444-1
Sequence 1, Application US/08466444
Patent No. 5776676

GENERAL INFORMATION:
APPLICANT: Prasad Ph.D., Gaddamannu L.
APPLICANT: Cooper M.D., Herbert L.
TITLE OF INVENTION: EPITHELIAL CELL SPECIFIC DIFFERENTIATION
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 W. Broadway, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,444

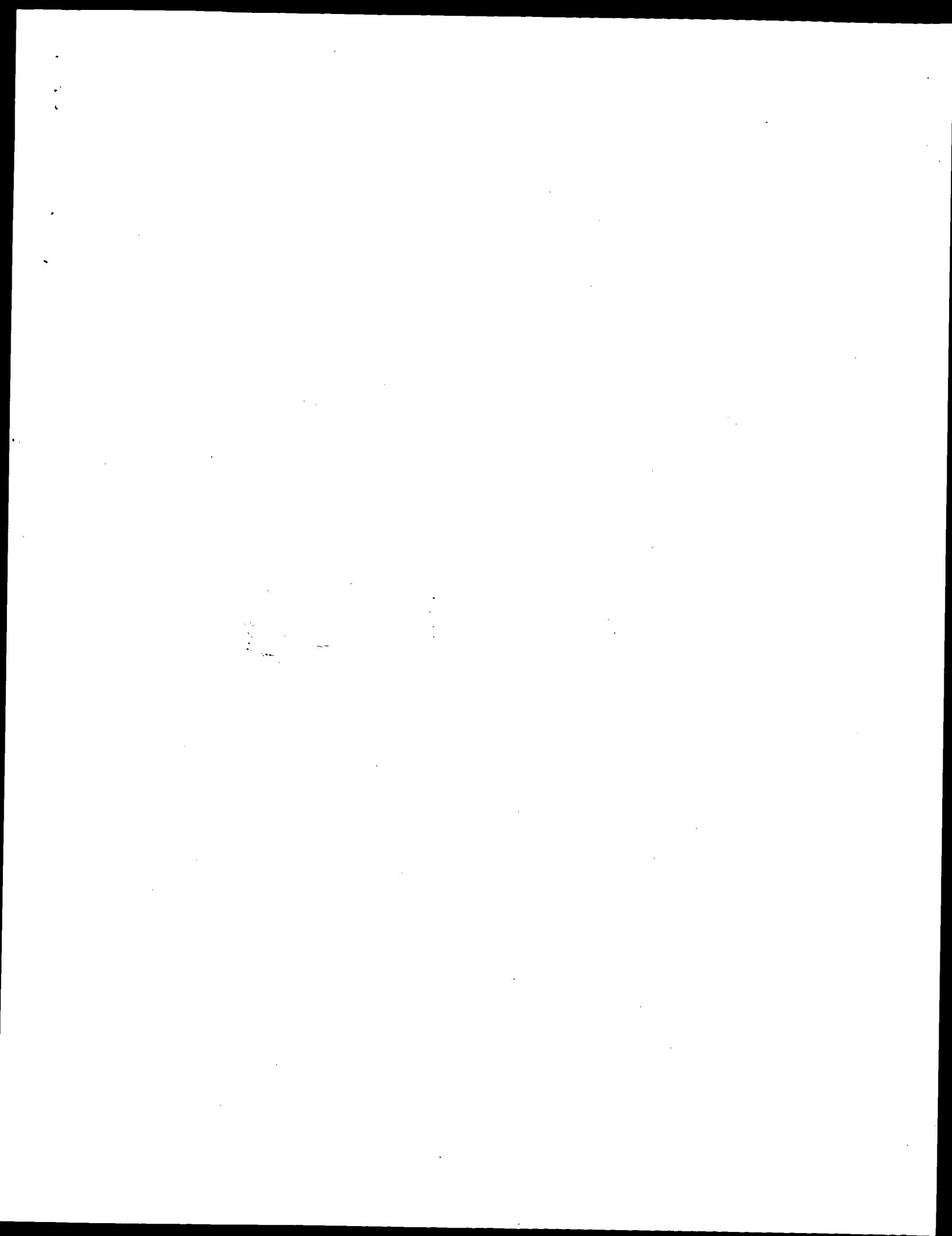
FILED DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/887,072
FILED DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kirkpatrick Ph.D., Anita M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH021.021A
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-466-444-1

Query Match
Best Local Similarity 8.7%; Score 32.8; DB 1; Length 1245;
Matches 161; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

QY 2 TGGGGTGGGGGACCTGCTCTATTCCTGATGTGACATTCGACGCGCTTTCA 61
DB 309 TGGGGCTGCTGGACACCCATCTATCAAGAGGCGCGGAGCCGAGAGCCGGGTCTTC 368
QY 62 AGTATGCTGATCCAGTCCACTGCGTCCCGCTCCGGGCTCCGGGTGACAGAGA 121
DB 369 ACCTGAAGATGAAGGGTGTACTTACCTGCTGCTGCGGAGGTGACACCGGTGACGACA 428
QY 122 AGAGATCGTGGGGCTTCAAGTGGGCTGATACCATGCGGACATTCAGCAAAAGTGT 181
DB 429 AGAAGGCGATCATTTGACTAGCCCGGTGACCTACCAAGAGGCGCATGTGACATCAGCAAGA 488
QY 182 CGGGGACATGCAAGAGCAAGGCTGCACTGTGATGTGTGGGGGGGGGGGCAATCC 241
DB 489 AG---GAGATGCCGCCACCAACCCCACTCCGCTGGGCTGCGCTGAACTTTTCCGT 545
QY 242 ACCAGAGTGAAGCAAGAAATTCACGTGTACGCTATTCATGAGGCTTATGCTGCC 301
DB 546 TCCACTAGAGATGCGCAACAGCCCGGAGAGGCGCATCTCTGCGCAAGACCACTTTCG 605
QY 302 AGCAGCCATTTCACCTGAGAAATCAAGCCAGTACCCGCGCTAGAGAGTCC 357
DB 606 ACGAGGCGATGCTGATCTGCAACCCCTCAGCGAGGACTCTTACAAAGACAGACCC 661

RESULT 14
US-09-210-748A-1
Sequence 1, Application US/09210748A
Patent No. 6335156

GENERAL INFORMATION:
APPLICANT: Hermeking, Heiko
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: 14-3-3 SIGMA ARREST THE CELL CYCLE
FILE REFERENCE: 1107.77810
CURRENT APPLICATION NUMBER: US/09/210,748A
CURRENT FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/069,416
PRIOR FILING DATE: 1997-12-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1320
TYPE: DNA
ORGANISM: Homo sapiens
US-09-210-748A-1



Result						Description	
No.	Score	Query	Match	Length	DB	ID	
1	339.8	90.6	422	9	US-09-918-995-4767		Sequence 4767, App
C	145	38.7	369	10	US-09-880-107-3508		Sequence 3508, App
	3	130.8	34.9	9	US-09-918-995-1676		Sequence 1676, App
	4	106.2	28.3	461	9	US-09-918-995-29191	Sequence 29191, App
C	5	89.4	23.6	330	10	US-09-880-107-2866	Sequence 2866, App
	6	39.6	10.8	2384	10	US-09-825-262-9705	Sequence 9705, App
7	38.2	10.2	891	9	US-10-166-081-41		Sequence 41, App
8	38.2	10.2	32559	9	US-10-166-081-1		Sequence 1, App
9	37.4	10.0	2436	10	US-09-815-242-7607		Sequence 7607, App
10	35.6	9.5	330	10	US-09-790-339-7		Sequence 7, App
11	34.6	9.2	1609	9	US-10-082-830-48		Sequence 48, App
12	33.8	9.0	536165	9	US-09-939-964-1		Sequence 1, App
13	33.4	8.9	594	9	US-10-123-155-10		Sequence 10, App
14	33	8.8	1119	9	US-10-123-155-352		Sequence 352, App
15	32.8	8.8	1185	10	US-09-887-576-784		Sequence 784, App
16	32.8	8.7	1330	10	US-09-939-581A-1		Sequence 1, App
17	32.8	8.7	1450	10	US-09-880-107-3741		Sequence 3741, App
18	32.8	8.7	2622	10	US-09-950-370-3		Sequence 3, App
19	32.8	8.7	2706	10	US-09-950-370-13		Sequence 13, App

ALIGNMENTS

348 TCACGACCAAGAAT

CTGCCACTGTGAGGGTGTGGCGGGGGCCGATCTCCACACGAG 240

249 TCAGACACAGAGATTCACTGTACGGCTATTTCATGGCTATGCTCTGCCAGCAGCC 308

NUMBER OF SEQ ID NOS: 38054

FILE REFERENCE: 44921-5028-WO

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

```

CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2866
LENGTH: 330
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142861A1 R45994
NAME/KEY: unsure
LOCATION: (1)..(330)
OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2866

Query Match          23.8%  Score 89.4;  DB 10;  Length 330;
Best Local Similarity 95.7%;  Pred. No. 3e-20;
Matches 90; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  282 CATGGCCTATGTCCTCCGACGACGCGCTTCACTGAGAAAATCAAGCAAGTACCC 341
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  285 CAAGGCTATGTCCTCTCCGACGACGCGCATTTCACTGAGAAAATCAAGCAAGTACCC 226
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY  342 CGACTACGAGTCACCTGGGCTTAACGACGCGTAC 375
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  225 CGACTACGAGTCACCTGGGCTTAACGAGGCTAC 192
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 6
US-09-815-242-9705
Sequence 9705, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9705
LENGTH: 2544
TYPE: DNA
ORGANISM: Salmonella typhi
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2544)

```

```

NAME/KEY: misc.feature
LOCATION: (1)...(2544)
OTHER INFORMATION: n = A,T,C or G
US-09-815-242-9705

Query Match          10.6%  Score 39.6;  DB 10;  Length 2544;
Best Local Similarity 50.8%;  Pred. No. 0.0032;
Matches 93; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY  31 GATGTGACATGACTCCGACGCGCTTTCAAGTATGTCGTGATCCGATCCACTCGGCT 90
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  691 GAGTTGAGAGTGCATTTATCTCTTCTACCGGATTCGGCCCGGCAAGCTGGCT 750
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  91 CCCGCTCCGGGGCTCCGGCTCCAGAGACAGAGATGTCGGCGCTACAAAGTGGCT 150
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  751 ACCCAGTGGGCTGTGTCGTGTCGCGACGCGCGGGGATGATTAAGACACCTGGCAG 810
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  151 GAGTACCATGGGACATCTACGACAAAGTGTGCGGACATGACAGACAGCTGGCAGC 210
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  811 ACCTCCGACCGGATATCTACGCTATCGGAGAGTGTGCGGAGTTGAGACACCGCTTAC 870
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  211 TGT 213
    ||
Db  871 GGT 873

RESULT 7
US-10-166-087-41
Sequence 41, Application US/10166087
Publication No. US2003007767A1
GENERAL INFORMATION:
APPLICANT: Ecopia Biosciences Inc.
APPLICANT: Farnet, Chris
APPLICANT: Staffa, Alfredo
APPLICANT: Zazopoulos, Emmanuel
TITLE OF INVENTION: Genes and proteins for the biosynthesis of anthramycin
FILE REFERENCE: 3014-205
CURRENT APPLICATION NUMBER: US/10/166,087
CURRENT FILING DATE: 2002-06-11
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn version 3.0
SEQ ID NO 41
LENGTH: 891
TYPE: DNA
ORGANISM: Streptomyces refuineus subspecies thermotolerans
US-10-166-087-41

Query Match          10.2%  Score 38.2;  DB 9;  Length 891;
Best Local Similarity 49.3%;  Pred. No. 0.0072;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY  84 CTCGGCTCCCGCTCCGGGCTCCGGCTGCGAGAGCAAGAGATGTCGGCGCTACAA 143
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  177 CGCGGCGCCGGGACACCGGGACCCCGGACCCCGCTGTGTCGACGCGCGGCA 236
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  144 GTGGGCTGATACATCGGACATCTACGACAAAGTGTGCGGACATGAGAAAGG 203
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  237 CTGGCAGAGTCCGCAATCGACGACGCTCTTGGCGGACGGAAGCGCTTGGCGACGG 296
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  204 CTGGGACTGTGAGTGTGTGGGCGGCGCATCTCCACACAGATGACAGCAAGAGAT 263
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  297 ATGGGCGTGTGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 356
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  264 TCACGTGATGCGCTATTCATGG 286
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  357 CGCTCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 379
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-10-166-087-1
Sequence 1, Application US/10166087
Publication No. US2003007767A1
GENERAL INFORMATION:

```

APPLICANT: Ecopia Biosciences Inc.
 APPLICANT: Farnel, Chris
 APPLICANT: Staifa, Alfredo
 APPLICANT: Zazopoulos, Emmanuel
 TITLE OF INVENTION: Genes and proteins for the biosynthesis of anthracyclin
 FILE REFERENCE: 3014-205
 CURRENT APPLICATION NUMBER: US/10/166,087
 CURRENT FILING DATE: 2002-06-11
 NUMBER OF SEQ ID NOS: 51
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 1
 LENGTH: 32539
 TYPE: DNA
 ORGANISM: Streptomyces refuineus subspecies thermotolerans
 US-10-166-087-1

Query Match 10.2%; Score 38.2; DB 9; Length 32539;
 Best Local Similarity 49.3%; Pred. No. 0.019;
 Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 84 CTGGGCTCCCGCTCCGGGCTCCGGCTGCAGAGACAGAGATGTCGGCGCTACAA 143
 DB 21902 CGCGGCGCCCGGACACACCGGACCCCGGACCCCGCTGCTGTCGACGGCGGCA 21961
 QY 144 GTGGGCTAGTACAGCGGACATCTACGACAAAGTGTGGGAGATGACAGAGAG 203
 DB 21962 CTGGCAGAGATCCGGCATCAGACAGCGCTGCTTCCGCGGACGACGCGCTGGCGGCG 22021
 QY 204 CTGGCAGTGTAGTGTCTGGGCGGCGGCGCATCTCCACCAAGTACAGAGAGAGAT 263
 DB 22022 ATCGGGTTCGTGGCGGCTACGAGGCTCGCGGCGGACGACGCTGCCGACATGAT 22081
 QY 264 TCACGGTACGGCTATTCATG 286
 DB 22082 CGCTCGGTGGCCCGGCGCTCG 22104

RESULT 9

US-09-815-242-7607
 Sequence 7607 Application US/09815242
 Patent No. US20020061569A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA 011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 7607
 LENGTH: 2436

TYPE: DNA
 ORGANISM: Klebsiella pneumoniae
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(2436)
 US-09-815-242-7607

Query Match 10.0%; Score 37.4; DB 10; Length 2436;
 Best Local Similarity 47.6%; Pred. No. 0.017;
 Matches 110; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 12 GGACCTCGCTCTCATCTCTGATGACATGACATCCGACGCGCTTCAGTATGCT 71
 DB 564 GCACCTCGCGGACGCTAGGACATGACATCTGATCTTCAACCGGATCG 623
 QY 72 GATCCGATCCACTGCTGCTCCCGCTCCGGGCTCCGCTCCGAGAGAGAGATCGT 131
 DB 624 CCGGCGGCTAAGCTGCGCCACCACTGCGGCTGCGGCTGCGGCGGCGCATCAT 683
 QY 132 GCGGCGCTACAAAGTGGGTGAGTACATCGGACATCTACAGCAAGTGTGGCGCAT 191
 DB 684 GGTCAACGATAGCTCCGACACTCCGATCCGATATCTATCTACATCGGGAATCGCAG 743
 QY 192 GCAGAGCAAGGCTGCGACTGTGAGTGTCTGGGCGGCGGCGCATCTCCCA 242
 DB 744 CTGGAACAACCGCGCTCTACGCTGTGTCGCCCGGCGCTATTAATGGCCCA 794

RESULT 10

US-09-790-399-7
 Sequence 7 Application US/09790399
 Patent No. US20020038000A1
 GENERAL INFORMATION:
 APPLICANT: Tuerk, Larry
 APPLICANT: Gold, Craig
 APPLICANT: Pridnow, David
 APPLICANT: Smith, Jonathan D.
 TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
 FILE REFERENCE: NEX02/C1-CON2
 CURRENT APPLICATION NUMBER: US/09/790,399
 CURRENT FILING DATE: 2001-02-22
 PRIOR APPLICATION NUMBER: 09/197,649
 PRIOR FILING DATE: 1998-11-23
 PRIOR APPLICATION NUMBER: 07/829,461
 PRIOR FILING DATE: 1992-01-31
 PRIOR APPLICATION NUMBER: 07/739,055
 PRIOR FILING DATE: 1991-08-01
 PRIOR APPLICATION NUMBER: 07/561,968
 PRIOR FILING DATE: 1990-08-02
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 7
 LENGTH: 390
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Sequence
 OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
 OTHER INFORMATION: fragments having NcoI restriction sites.
 US-09-790-399-7

Query Match 9.5%; Score 35.6; DB 10; Length 390;
 Best Local Similarity 43.5%; Pred. No. 0.042;
 Matches 161; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 3 GGGGTTGGGGGACCTGCTCTCATCTTCTATGAGCATGCACTCCGAGCGGCTTCAA 62
 DB 2 GGCATGAG 61
 QY 63 GATGTGCTGATCCAGTCCACTGCTCCCGCTCCGGGCTCCGGGCTCCGAGAGAGCA 122
 DB 62 CGAG 121

QY 70 CTGATCCGAGTCCAGTCCGCTCCCGGCTCCGAGGCTCCGAGAGAGAGATC 129
DB 367 PCRNASARRRRSSSSSSASRTSSSSSSSSRRGGGYTSGHANSRFSRSR 426
QY 130 GTGGCGGCTCAAGTGGCTGAGTACATCGGACATCAAGCAAAAGTGTGGGCGAC 189
DB 427 SRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRRSRR 486
QY 190 ATGCAAGAGCAAGCTCGGATGTGCTGTGGGCGGCGGCGCC 234
DB 487 RHSSSRSSSSSLSPSRSLTSSRSHSPSPSSRSRSRSS 531

RESULT 14

US-10-123-155-352
Sequence 352, Application US/10123155
Publication No. US20030068794A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gunney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 350
SEQ ID NO 352
LENGTH: 1119
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-155-352

Query Match 8.8%; Score 33; DB 9; Length 1119;
Best Local Similarity 12.1%; Pred. No. 0.41;
Matches 32; Conservative 82; Mismatches 151; Indels 0; Gaps 0;

QY 34 GTGACATCCAGTCCGAGGCTTTCAGATGTGCTGATCCGAGTCCAGTCCGCTCC 93
DB 821 GTSLVWVYIYHTRRRNEDSITNTDETNPADIPYLSQGLADQDYVSESSSH 880
QY 94 CGCTCCGAGGCTCCGCTGAGAGAGAGATCGTGGCGGCTCAAGTGGGCTGAG 153
DB 881 QFVTSAGAGFLPQHDSSGTHIDNSSEADYEATDLEFLPGLASTPMLKGVISDP 940
QY 154 TACATCGGAGCATCTAGACAAAGTGTGGGCGAGATGAGAGCAAGAGCTGGGACT 213
DB 941 FETVHTGSPDPRTVLMHDYEPYIKKKECPGSHPESECESEFNSIMPSHVKLNT 1000
QY 214 GAGTGTCTGGGCGGCGGAGATCTCCACACAGAGTCAAGAGCAAGAGATTCAGCTAC 273
DB 1001 SYSNHEGPMKNTCLNKSISLDFSANPPASVANSNSFMGTGKALRPDLAYSSFGOPS 1060
QY 274 GGCTATCCAGTGGCTATGGTCTG 298
DB 1061 DQPRATYLRKHAHSSPDLDGSEEDG 1085

RESULT 15

US-09-887-576-784
Sequence 784, Application US/09887576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, B.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001US1
CURRENT APPLICATION NUMBER: US/09/887,576
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/213,848
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/214,087
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/258,692
PRIOR FILING DATE: 2000-12-29
SOFTWARE: FastSeq for Windows Version 4.0
NUMBER OF SEQ ID NOS: 875
SEQ ID NO 784
LENGTH: 1185
TYPE: DNA
ORGANISM: Oryza sativa
US-09-887-576-784

Query Match 8.7%; Score 32.8; DB 10; Length 1185;
Best Local Similarity 44.3%; Pred. No. 0.49;
Matches 133; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 13 GACCTCGCTCATTCCTGATGTGACATGACATCCGAGCGGTCTCAAGTATGTGCTG 72
DB 229 GGCATGGCTGTGCTGTCGACGACGCGGCTCGACGCGGCTGCAAGGTGCTGCTC 288
QY 73 ATCCGAGTCCAGTCCGCTCCCGGCTCCGCGGCTCCGCGGCTCCGAGAGAGATGCTG 132
DB 289 AACATGACGACGACATGTCGCGACATCGGAGGAGGCTGACGCGCATCTTCAACCAAGCC 348
QY 133 CGGAGTACAGTGGCTGAGTACATGAGTCCGAGCATCTACAGCAAAAGTGTGGGCGACATG 192
DB 349 CCCGAGAGATGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 408
QY 193 CAGAGCAAGGCTGCGATGTGATGTGTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 252
DB 409 CCCGAGTATGCTCCCTCAGCCAGCTGCTCGCCACCAAGCTGGGCGGCGGCGGCGGCGGCGG 468
QY 253 GACCAAGATTCAGTGTGAGGCTATTCATGGGCTATGGCTGCTGCTGCTGCTGCTGCTGCT 312
DB 469 GTCCGCAAGAGGCACTGCGCTGTGCTCAGGCGCCGAGGCGGCAAGAGCAAGTACCGCTT 528

Search completed: May 13, 2003, 20:30:16
Job time : 565 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 7, 2003, 19:04:05 ; Search time 30.3899 Seconds

(without alignments)
378.521 Million cell updates/sec

Title: US-09-914-831-2

Perfect score: 674
Sequence: 1 MAVADIALIPVDIDSDGVF.....TEKIKKPYDEYTWANDGY 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Published_Applications_AA.*
2: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	11.6	351	10 US-09-864-761-33332	Sequence 33332, A
2	73	10.8	149	9 US-09-991-496-86	Sequence 86, App1
3	73	10.8	149	10 US-09-874-923-86	Sequence 86, App1
4	71	10.5	224	9 US-09-738-626-5900	Sequence 5900, App
5	71	10.5	392	10 US-09-945-825-8	Sequence 8, App1
6	70.5	10.5	484	9 US-10-092-063-27	Sequence 27, App1
7	68.5	10.2	5701	10 US-09-864-761-37319	Sequence 37319, A
8	67	9.9	344	9 US-10-078-770-100	Sequence 100, App
9	66	9.8	263	10 US-09-815-242-11078	Sequence 11078, A
10	65.5	9.7	159	9 US-09-992-598-361	Sequence 361, App
11	65.5	9.7	159	9 US-09-989-735-361	Sequence 361, App
12	65.5	9.7	159	9 US-09-989-735-361	Sequence 361, App
13	65.5	9.7	159	9 US-09-990-444-361	Sequence 361, App
14	65.5	9.7	159	9 US-09-989-730-361	Sequence 361, App
15	65.5	9.7	159	9 US-09-990-436-361	Sequence 361, App
16	65.5	9.7	159	9 US-09-991-181-361	Sequence 361, App
17	65.5	9.7	159	9 US-09-993-687-361	Sequence 361, App
18	65.5	9.7	159	9 US-09-989-734-361	Sequence 361, App
19	65.5	9.7	159	9 US-09-997-653-361	Sequence 361, App

20	65.5	9.7	159	9 US-10-174-590-240	Sequence 240, App
21	65.5	9.7	159	9 US-10-176-758-240	Sequence 240, App
22	65.5	9.7	159	9 US-10-175-737-240	Sequence 240, App
23	65.5	9.7	159	9 US-09-993-667-361	Sequence 361, App
24	65.5	9.7	159	9 US-10-173-706-240	Sequence 240, App
25	65.5	9.7	159	9 US-10-175-738-240	Sequence 240, App
26	65.5	9.7	159	9 US-10-175-752-240	Sequence 240, App
27	65.5	9.7	159	9 US-10-176-482-240	Sequence 240, App
28	65.5	9.7	159	9 US-10-176-757-240	Sequence 240, App
29	65.5	9.7	159	9 US-10-176-913-240	Sequence 240, App
30	65.5	9.7	159	9 US-10-180-552-240	Sequence 240, App
31	65.5	9.7	159	9 US-10-180-557-240	Sequence 240, App
32	65.5	9.7	159	9 US-09-990-438-361	Sequence 361, App
33	65.5	9.7	159	9 US-09-990-362-361	Sequence 361, App
34	65.5	9.7	159	9 US-09-997-428-361	Sequence 361, App
35	65.5	9.7	159	9 US-09-997-666-361	Sequence 361, App
36	65.5	9.7	159	9 US-10-173-700-240	Sequence 240, App
37	65.5	9.7	159	9 US-10-174-572-240	Sequence 240, App
38	65.5	9.7	159	9 US-10-174-579-240	Sequence 240, App
39	65.5	9.7	159	9 US-10-174-582-240	Sequence 240, App
40	65.5	9.7	159	9 US-10-174-588-240	Sequence 240, App
41	65.5	9.7	159	9 US-10-175-739-240	Sequence 240, App
42	65.5	9.7	159	9 US-10-175-740-240	Sequence 240, App
43	65.5	9.7	159	9 US-10-175-743-240	Sequence 240, App
44	65.5	9.7	159	9 US-10-176-488-240	Sequence 240, App
45	65.5	9.7	159	9 US-10-176-492-240	Sequence 240, App

ALIGNMENTS

RESULT 1
US-09-864-761-33332
Sequence 33332, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomics-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5900
LENGTH: 224
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5900

Query Match 10.5%; Score 71; DB 9; Length 224;
Best Local Similarity 25.2%; Pred. No. 3;
Matches 29; Conservative 15; Mismatches 33; Indels 38; Gaps 8;

QY 23 VLIRVHAPRSGAPAAESKEIVRGYKMAEYHADIDYKSGDMKQCCD-----70
DB 116 MLIRLTADAERKEVEREGE---SFHYGDYVA--YSKICTHI---GCPSTLYEAQTNRII 167
QY 71 CECILGGGRISHOSQDKIRHYGYSMAVGPQAHAISTEKIKAKYPPDEVYVWANDGY 125
DB 168 CPC-----HOSQFDALH-YG-KPVFGPAARAL-----GCPITVDEEGY 204

RESULT 5
US-09-945-825-8
Sequence 8, Application US/09945825
Patent No. US2002010669A1
GENERAL INFORMATION:
APPLICANT: NOBUHITO, SONE
TITLE OF INVENTION: Respiratory Chain Enzyme Genes of Coryneform Bacteria
FILE REFERENCE: 213639US-8222-10-0
CURRENT APPLICATION NUMBER: US/09/945,825
CURRENT FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: JP 2000-270283
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 392
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-945-825-8

Query Match 10.5%; Score 71; DB 10; Length 392;
Best Local Similarity 25.2%; Pred. No. 6.2;
Matches 29; Conservative 15; Mismatches 33; Indels 38; Gaps 8;

QY 23 VLIRVHAPRSGAPAAESKEIVRGYKMAEYHADIDYKSGDMKQCCD-----70
DB 284 MLIRLTADAERKEVEREGE---SFHYGDYVA--YSKICTHI---GCPSTLYEAQTNRII 335
QY 71 CECILGGGRISHOSQDKIRHYGYSMAVGPQAHAISTEKIKAKYPPDEVYVWANDGY 125
DB 336 CPC-----HOSQFDALH-YG-KPVFGPAARAL-----GCPITVDEEGY 372

RESULT 6
US-10-092-063-27
Sequence 27, Application US/10092063
Patent No. US20020173005A1
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mleto, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES

FILE REFERENCE: 28110/35908
CURRENT APPLICATION NUMBER: US/10/092,063
CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-063-27

Query Match 10.5%; Score 70.5; DB 9; Length 484;
Best Local Similarity 29.6%; Pred. No. 9.2;
Matches 32; Conservative 12; Mismatches 31; Indels 33; Gaps 7;

QY 34 GAPAAESKEIVR-----GYKMAEYHADIDYKSGDMKQCCDCEICGGRIS-----H 81
DB 312 GOPANDGKEVYPCLESPFKGEWEHETVYRVSG--OKAAASLHEILCAARSEVLONRVH 369
QY 82 OSQDKIRHY-----YGSMAVGPQAHAISTEK-----IKAKY 113
DB 370 KIEEVK-HVDYFATSYTYDIAAGV--LIDAEKGSILVGDPEIAKAY 414

RESULT 7
US-09-864-761-37319
Sequence 37319, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmicha-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668

```

      1  PRIOR APPLICATION NUMBER: 60/143,409
      2  PRIOR FILING DATE: 1999-07-12
      3  NUMBER OF SEQ ID NOS: 196
      4  SOFTWARE: Microsoft Office 97
      5  SEQ ID NO 100
      6  LENGTH: 344
      7  TYPE: PRT
      8  ORGANISM: Glycine max
      9  US-10-078-770-100
      10
      11 Query Match
      12 Best Local Similarity 21.8%; Pred. No. 15;
      13 Matches 22; Conservative 23; Mismatches 44; Indels 12; Gaps 2
      14
      15
      16
      17
      18
      19
      20
      21 KYILRVHSPNSGAPAAES-KEIVRGYKAEYHADLYDKVSDMKOGCCDECLGGGR1 79
      22 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      23
      24 DB 131 RYRLKKNNEVSGMAVAQSPRRILYLDRPLQLNKDDYDRVCQIPQKGANFRLDG--- 187
      25 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      26
      27
      28
      29
      30
      31
      32
      33
      34
      35
      36
      37
      38
      39
      40
      41
      42
      43
      44
      45
      46
      47
      48
      49
      50
      51
      52
      53
      54
      55
      56
      57
      58
      59
      60
      61
      62
      63
      64
      65
      66
      67
      68
      69
      70
      71
      72
      73
      74
      75
      76
      77
      78
      79
      80
      81
      82
      83
      84
      85
      86
      87
      88
      89
      90
      91
      92
      93
      94
      95
      96
      97
      98
      99
      100
      101
      102
      103
      104
      105
      106
      107
      108
      109
      110
      111
      112
      113
      114
      115
      116
      117
      118
      119
      120
      121
      122
      123
      124
      125
      126
      127
      128
      129
      130
      131
      132
      133
      134
      135
      136
      137
      138
      139
      140
      141
      142
      143
      144
      145
      146
      147
      148
      149
      150
      151
      152
      153
      154
      155
      156
      157
      158
      159
      160
      161
      162
      163
      164
      165
      166
      167
      168
      169
      170
      171
      172
      173
      174
      175
      176
      177
      178
      179
      180
      181
      182
      183
      184
      185
      186
      187
      188
      189
      190
      191
      192
      193
      194
      195
      196
      197
      198
      199
      200
      201
      202
      203
      204
      205
      206
      207
      208
      209
      210
      211
      212
      213
      214
      215
      216
      217
      218
      219
      220
      221
      222
      223
      224
      225
      226
      227
      228
      229
      230
      231
      232
      233
      234
      235
      236
      237
      238
      239
      240
      241
      242
      243
      244
      245
      246
      247
      248
      249
      250
      251
      252
      253
      254
      255
      256
      257
      258
      259
      260
      261
      262
      263
      264
      265
      266
      267
      268
      269
      270
      271
      272
      273
      274
      275
      276
      277
      278
      279
      280
      281
      282
      283
      284
      285
      286
      287
      288
      289
      290
      291
      292
      293
      294
      295
      296
      297
      298
      299
      300
      301
      302
      303
      304
      305
      306
      307
      308
      309
      310
      311
      312
      313
      314
      315
      316
      317
      318
      319
      320
      321
      322
      323
      324
      325
      326
      327
      328
      329
      330
      331
      332
      333
      334
      335
      336
      337
      338
      339
      340
      341
      342
      343
      344
      345
      346
      347
      348
      349
      350
      351
      352
      353
      354
      355
      356
      357
      358
      359
      360
      361
      362
      363
      364
      365
      366
      367
      368
      369
      370
      371
      372
      373
      374
      375
      376
      377
      378
      379
      380
      381
      382
      383
      384
      385
      386
      387
      388
      389
      390
      391
      392
      393
      394
      395
      396
      397
      398
      399
      400
      401
      402
      403
      404
      405
      406
      407
      408
      409
      410
      411
      412
      413
      414
      415
      416
      417
      418
      419
      420
      421
      422
      423
      424
      425
      426
      427
      428
      429
      430
      431
      432
      433
      434
      435
      436
      437
      438
      439
      440
      441
      442
      443
      444
      445
      446
      447
      448
      449
      450
      451
      452
      453
      454
      455
      456
      457
      458
      459
      460
      461
      462
      463
      464
      465
      466
      467
      468
      469
      470
      471
      472
      473
      474
      475
      476
      477
      478
      479
      480
      481
      482
      483
      484
      485
      486
      487
      488
      489
      490
      491
      492
      493
      494
      495
      496
      497
      498
      499
      500
      501
      502
      503
      504
      505
      506
      507
      508
      509
      510
      511
      512
      513
      514
      515
      516
      517
      518
      519
      520
      521
      522
      523
      524
      525
      526
      527
      528
      529
      530
      531
      532
      533
      534
      535
      536
      537
      538
      539
      540
      541
      542
      543
      544
      545
      546
      547
      548
      549
      550
      551
      552
      553
      554
      555
      556
      557
      558
      559
      560
      561
      562
      563
      564
      565
      566
      567
      568
      569
      570
      571
      572
      573
      574
      575
      576
      577
      578
      579
      580
      581
      582
      583
      584
      585
      586
      587
      588
      589
      590
      591
      592
      593
      594
      595
      596
      597
      598
      599
      600
      601
      602
      603
      604
      605
      606
      607
      608
      609
      610
      611
      612
      613
      614
      615
      616
      617
      618
      619
      620
      621
      622
      623
      624
      625
      626
      627
      628
      629
      630
      631
      632
      633
      634
      635
      636
      637
      638
      639
      640
      641
      642
      643
      644
      645
      646
      647
      648
      649
      650
      651
      652
      653
      654

```

DB 73 QNVDINVKAFGTG-----DISTEMLKDFGAKY 99

RESULT 10
US-09-992-598-361
Sequence 361, Application US/09992598
Patent No. US20020160384A1

GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same

FILE REFERENCE: P2730PIC20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254

;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090662
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090663
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 9.7%; Score 65.5; DB 9; Length 159;
Best Local Similarity 23.7%; Pred. No. 8.3;
Matches 23; Conservative 19; Mismatches 32; Indels 23; Gaps 4;

OY 8 LIPDDIDSDGVKTYLVINSHAPSPGAPAEKSEIVRGYKWAEXHADIDKVSQDMOKQ 67
DB 24 LIPNTVLLLELLSKYOHNSHRSVRRAIPREDKEELM-----LHNTLRQOVPO 73
OY 68 GCDCECL--GGGRIS-HQSODKKTTHVGYSAATGPA 100
DB 74 ASNMEYWSAGSGRGWRH-----GWSIGHQPA 101

RESULT 11
US-09-989-293A-361
Sequence 361, Application US/09989293A
Patent No. US20020177164A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C66
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: US/09/989, 293A
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1993-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1993-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 9.7%; Score 65.5; DB 9; Length 159;
Best Local Similarity 23.7%; Pred. No. 8.3;
Matches 23; Conservative 19; Mismatches 32; Indels 23; Gaps 4;

QY 8 LIPVDIDSDGVFKYVILIRHSAPRSAGAPAESEKTEIVRGYKMAEYHADIDYKVSQDMQK 67
DB 24 LIPVTLLEELLSKYQHNEHSRVRKRAIPREDKEELM-----LHNTKRGQVQ 73
QY 68 GCDCECL--GGGRIS-HQSQDKKIHVYGSMAVGA 100
DB 74 ASNMEYVVSAGSRGRGWHR-----GWLGHQPA 101

RESULT 12
US-09-989-735-361
; Sequence 361, Application US/09989735

Publication No. US20020193299A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gunney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C61
CURRENT APPLICATION NUMBER: US/09/989,735
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429

```

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          9.7%: Score 65.5; DB 9; Length 159;
Best Local Similarity 23.7%: Pred. No. 8.3;
Matches 23; Conservative 19; Mismatches 32; Indels 23; Gaps 4;

QY 8 LIPVDIDSDVFKYVLTFRVSHAPPSGAPAESKEIYVRGYKWAYEYHADIYDKVSGDMQKQ 67
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 24 LIPVNTLEELLISKYQHSHESRVRRAIRPRDDKEITLM-----LHKKLNGQVQPQ 73

QY 68 GCDCECL--GGGRIS-HOSQDKIKHYGYSMAYGPA 100
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 74 ASNMEYVVSAGSGRRGMR-----GWSLGHQPA 101

RESULT 13
; Sequence 361, Application US/09990444
; Publication No. US20020193300A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C19
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
```

PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 9.7%; Score 65.5; DB 9; Length 159;
Best Local Similarity 23.7%; Pred. No. 8.3; Mismatches 23; Conservative 19; Indels 23; Gaps 4;

OY 8 LIPDVDDSDGVFVYILRVHAPRSCAPAESEKIVRGYMAEYHADIYDKVSGMQRQ 67
DB 24 LIPVTLLEELLSKYQHNSHRSVRRAIPREDKEELM-----LHNTLRGVQPO 73
OY 68 GDCDECU---GGGRIS-HOSODKKIHVYGSMAVGA 100
DB 74 ASNMEYVWSAGSGRGRWHR-----GWGLGHOPA 101

RESULT 14
US-09-989-730-361
Sequence 361, Application US/09989730
Publication No. US20020197674A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc

Wed May 14 09:22:06 2003

us-09-914-831-2.rapb

Page 11

APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavln, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PLC69
CURRENT APPLICATION NUMBER: US/09/989,730
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444

us-09-914-831-2 ranch

1 PRIOR FILING DATE: 1998-06-24
 2 PRIOR APPLICATION NUMBER: 60/090445
 3 PRIOR FILING DATE: 1998-06-24
 4 PRIOR APPLICATION NUMBER: 60/090472
 5 PRIOR FILING DATE: 1998-06-24
 6 PRIOR APPLICATION NUMBER: 60/090535
 7 PRIOR FILING DATE: 1998-06-24
 8 PRIOR APPLICATION NUMBER: 60/090540
 9 PRIOR FILING DATE: 1998-06-24
 10 PRIOR APPLICATION NUMBER: 60/090542
 11 PRIOR FILING DATE: 1998-06-24
 12 PRIOR APPLICATION NUMBER: 60/090557
 13 PRIOR FILING DATE: 1998-06-24
 14 PRIOR APPLICATION NUMBER: 60/090676
 15 PRIOR FILING DATE: 1998-06-25
 16 PRIOR APPLICATION NUMBER: 60/090678
 17 PRIOR FILING DATE: 1998-06-25
 18 PRIOR APPLICATION NUMBER: 60/090690
 19 PRIOR FILING DATE: 1998-06-25
 20 PRIOR APPLICATION NUMBER: 60/090694
 21 PRIOR FILING DATE: 1998-06-25
 22 PRIOR APPLICATION NUMBER: 60/090695
 23 PRIOR FILING DATE: 1998-06-25
 24 PRIOR APPLICATION NUMBER: 60/090696
 25 PRIOR FILING DATE: 1998-06-25
 26 PRIOR APPLICATION NUMBER: 60/090862
 27 PRIOR FILING DATE: 1998-06-26
 28 PRIOR APPLICATION NUMBER: 60/090863
 29 PRIOR FILING DATE: 1998-06-26
 30 PRIOR APPLICATION NUMBER: 60/091360
 31 PRIOR FILING DATE: 1998-07-01
 32 PRIOR APPLICATION NUMBER: 60/091478
 33 PRIOR FILING DATE: 1998-07-02
 34 PRIOR APPLICATION NUMBER: 60/091544
 35 PRIOR FILING DATE: 1998-07-01
 36 PRIOR APPLICATION NUMBER: 60/091519
 37 PRIOR FILING DATE: 1998-07-02
 38 PRIOR APPLICATION NUMBER: 60/091626
 39 PRIOR FILING DATE: 1998-07-02
 40 PRIOR APPLICATION NUMBER: 60/091653
 41 PRIOR FILING DATE: 1998-07-02
 42 PRIOR APPLICATION NUMBER: 60/091978
 43 PRIOR FILING DATE: 1998-07-07
 44 PRIOR APPLICATION NUMBER: 60/091982
 45 PRIOR FILING DATE: 1998-07-07
 46 PRIOR APPLICATION NUMBER: 60/092182
 47 PRIOR FILING DATE: 1998-07-09

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavich, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730F01C14
CURRENT APPLICATION NUMBER: US/09/990,436
PRIOR APPLICATION NUMBER: 2001-11-14
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/0649787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217

PRIOR FILING DATE: 1998-06-05	PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09	PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/0889105
PRIOR FILING DATE: 1998-06-12	PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089633
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090451

1	PRIOR APPLICATION NUMBER: 60/0904412
2	PRIOR FILING DATE: 1998-06-24
3	PRIOR APPLICATION NUMBER: 60/09050535
4	PRIOR FILING DATE: 1998-06-24
5	PRIOR APPLICATION NUMBER: 60/0905404
6	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/0905042
8	PRIOR FILING DATE: 1998-06-24
9	PRIOR APPLICATION NUMBER: 60/090557
10	PRIOR FILING DATE: 1998-06-24
11	PRIOR APPLICATION NUMBER: 60/090676
12	PRIOR FILING DATE: 1998-06-25
13	PRIOR APPLICATION NUMBER: 60/090678
14	PRIOR FILING DATE: 1998-06-25
15	PRIOR APPLICATION NUMBER: 60/090690
16	PRIOR FILING DATE: 1998-06-25
17	PRIOR APPLICATION NUMBER: 60/090695
18	PRIOR FILING DATE: 1998-06-25
19	PRIOR APPLICATION NUMBER: 60/090696
20	PRIOR FILING DATE: 1998-06-25
21	PRIOR APPLICATION NUMBER: 60/090699
22	PRIOR FILING DATE: 1998-06-25
23	PRIOR APPLICATION NUMBER: 60/090866
24	PRIOR FILING DATE: 1998-06-26
25	PRIOR APPLICATION NUMBER: 60/090865
26	PRIOR FILING DATE: 1998-06-26
27	PRIOR APPLICATION NUMBER: 60/091360
28	PRIOR FILING DATE: 1998-07-01
29	PRIOR APPLICATION NUMBER: 60/091478
30	PRIOR FILING DATE: 1998-07-02
31	PRIOR APPLICATION NUMBER: 60/091544
32	PRIOR FILING DATE: 1998-07-01
33	PRIOR APPLICATION NUMBER: 60/091519
34	PRIOR FILING DATE: 1998-07-02
35	PRIOR APPLICATION NUMBER: 60/091628
36	PRIOR FILING DATE: 1998-07-02
37	PRIOR APPLICATION NUMBER: 60/091633
38	PRIOR FILING DATE: 1998-07-02
39	PRIOR APPLICATION NUMBER: 60/091978
40	PRIOR FILING DATE: 1998-07-07
41	PRIOR APPLICATION NUMBER: 60/091983
42	PRIOR FILING DATE: 1998-07-07
43	PRIOR APPLICATION NUMBER: 60/092183
44	PRIOR FILING DATE: 1998-07-09

[illegible]

Search completed: May 7, 2003, 19:18:46
Job time : 33.3899 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:25 ; Search time 33.2569 Seconds

(without alignments)
110.590 Million cell updates/sec

Title: US-09-914-831-2

Perfect score: 674
Sequence: 1 MAVADIALIPVDIDSDGVF.....TEKIKAKYDYEWANDGY 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	11.6	445	2	US-08-935-450-4
2	76.5	11.4	388	2	US-08-382-505-2
3	73	10.8	149	4	US-09-183-861-86
4	73	10.8	149	4	US-09-022-765-86
5	70.5	10.5	456	4	US-09-240-639-2
6	70.5	10.5	471	4	US-09-608-285A-60
7	70.5	10.5	484	4	US-09-370-265-27
8	70.5	10.5	484	4	US-09-230-637-30
9	67.5	10.0	333	4	US-07-772-087-6
10	65.5	9.7	381	1	US-08-173-508-12
11	65.5	9.7	381	2	US-08-265-310-12
12	65.5	9.7	381	3	US-08-951-742-12
13	65.5	9.7	382	1	US-08-460-343B-2
14	65.5	9.7	382	1	US-08-398-028B-2
15	65.5	9.7	382	1	US-08-504-265B-2
16	65.5	9.7	382	4	US-09-255-502-2
17	65.5	9.7	382	4	US-09-178-155-2
18	65.5	9.7	382	4	US-09-445-472-31
19	65.5	9.7	382	4	US-09-445-472-31
20	65	9.6	380	6	RE34606-2
21	64.5	9.6	382	1	US-08-049-282B-2
22	64.5	9.6	382	2	US-08-537-715-2
23	64.5	9.6	382	5	PCT-US94-04173-2
24	64.5	9.6	383	2	US-08-537-715-4
25	64.5	9.6	383	2	PCT-US94-04173-4
26	64	9.5	228	3	US-09-286-690-12
27	63.5	9.4	1447	4	US-09-376-330-17

28	62.5	9.3	464	4	US-09-426-072-2	Sequence 2, Appli
29	61.5	9.1	197	4	US-08-952-445-26	Sequence 26, Appl
30	61.5	9.1	281	4	US-08-952-445-28	Sequence 28, Appl
31	61.5	9.1	352	1	US-07-923-260A-2	Sequence 2, Appli
32	61.5	9.1	286	2	US-08-826-267-2	Sequence 2, Appli
33	61	9.1	376	2	US-08-485-449-5	Sequence 5, Appli
34	61	9.1	415	4	US-08-795-430-11	Sequence 11, Appl
35	61	9.1	415	4	US-09-352-700-11	Sequence 11, Appl
36	61	9.1	415	4	US-08-601-132-41	Sequence 41, Appl
37	61	9.1	1447	4	US-09-041-886-25	Sequence 25, Appl
38	61	9.1	1447	5	PCT-US94-05277-2	Sequence 2, Appli
39	60.5	9.0	352	1	US-07-923-260A-3	Sequence 3, Appli
40	60.5	9.0	368	2	US-08-903-624-3	Sequence 9, Appli
41	60.5	9.0	368	4	US-08-973-914-9	Sequence 9, Appli
42	60.5	9.0	381	1	US-07-772-087-5	Sequence 5, Appli
43	60.5	9.0	381	1	US-08-322-965-2	Sequence 2, Appli
44	60.5	9.0	382	6	5472855-2	Patent No. 5472855
45	60	8.9	384	4	US-09-552-322-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-935-450-4
; Sequence 4, Application US/08935450
; Patent No. 5977311
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; TITLE OF INVENTION: 53BP2 COMPLEXES
; FILE REFERENCE: 7934-054
; CURRENT APPLICATION NUMBER: US/08/935,450
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-935-450-4

Query Match      11.6% Score 78; DB 2; Length 445;
Best Local Similarity 25.7%; Pred. No. 0.13;
Matches 19; Conservative 13; Mismatches 28; Indels 14; Gaps 2;

OY 46 GYNAEYH---ADYKVSQDMKQCGDCCLGGGRISHOSQDKIHVGYSMAYGPAQ 101
    |||:| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB 98 GNNMAKGYTGCAELVSDVLDYVRKESRSCDCLQGFOLTH-----SLGGTSGM 147
    : ||: |||

OY 102 HAISTEKIKAKYPD 115
    : ||: |||
DB 148 GTLISKIREXPYD 161

RESULT 2
US-08-382-505-2
; Sequence 2, Application US/08382505
; Patent No. 5962765
; GENERAL INFORMATION:
; APPLICANT: St. Jeager, Raymond J.
; APPLICANT: Roberts, Donald W.
; APPLICANT: Staples, Richard C.
; TITLE OF INVENTION: Molecular Cloning of a Complimentary DNA
; TITLE OF INVENTION: Sequence Encoding a Cuticle Degrading Protease Produced By
; Patent No. 5962765
; TITLE OF INVENTION: Entomopathogenic Fungi
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BARNARD, BROWN & MICHAELS
; STREET: 306 East State Street, Suite 220
; CITY: Ithaca
```

STATE: NY
COUNTRY: USA
ZIP: 14850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/382,505
FILING DATE: 02-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/739,645
FILING DATE: 02-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A.
REGISTRATION NUMBER: 34,390
REFERENCE/DOCKET NUMBER: BTI-15CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 607-273-1711
TELEFAX: 607-273-2609
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-382-505-2

Query Match 11.4%; Score 76.5; DB 2; Length 388;
Best Local Similarity 26.7%; Pred. No. 0.17;
Matches 23; Conservative 20; Mismatches 38; Indels 5; Gaps 3;

QY 33 SGAPASKEIVAGCYKAEF-HADIDKVSQDMQKGCCEICGGGRISHOSQDKIHVY 91
DB 78 AGSLTKELMREHPVDIERDAVWRISGITGSGAP--WGLGRISHRSKSTTYR 134
QY 92 GYSWAYGPAQHAISTEKIRAKYPDY 117
DB 135 DDSAGQCTCYIIDT-CIEASHPEFE 159

RESULT 3
US-09-183-861-86
Sequence 86, Application US/09183861
Patent No. 6365165
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, David C.
APPLICANT: Skelky, Yasir A.W.
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,861
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/022,765

FILING DATE: 12-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.420C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-183-861-86

Query Match 10.8%; Score 73; DB 4; Length 149;
Best Local Similarity 28.4%; Pred. No. 0.13;
Matches 21; Conservative 11; Mismatches 28; Indels 14; Gaps 3;

QY 46 GYKAEYH---ADIDKVSQDMQKGCCEICGGGRISHOSQDKIHVYGYSWAYGPAQ 101
DB 53 GNNMAKGYTSCALIDVLYCKRKREASCCLQGFOLSHS-----LGGTSGMGRT 104
QY 102 HAISTEKIRAKYPD 115
DB 105 LLISX--LRXEPD 116

RESULT 4
US-09-022-765-86
Sequence 86, Application US/09022765
Patent No. 6375955
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, David C.
APPLICANT: Skelky, Yasir A.W.
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,765
FILING DATE: 12-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.420C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-022-765-86

Query Match 10.8%; Score 73; DB 4; Length 149;

```
Best Local Similarity 28.4%; Pred. No. 0.13;
Matches 21; Conservative 11; Mismatches 28; Indels 14; Gaps 3;
QY 46 GYKMAEYH---ADIDYKSGDMKQDCDCEGCRISHQSDKTIHYGYMAYGPAQ 101
Db 53 GNNMAKHYTEGALIDSVLDVCKEAKESDCDGLGFLSHS-----LGGGTGSGMGT 104
QY 102 HAISTEKIKAKYPD 115
Db 105 LLISX--LRXEYPD 116

RESULT 5
US-09-240-639-2
; Sequence 2, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischauf, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-240-639-2

Query Match 10.5%; Score 70.5; DB 4; Length 456;
Best Local Similarity 29.6%; Pred. No. 1.3;
Matches 32; Conservative 12; Mismatches 31; Indels 33; Gaps 7;
QY 34 GAPAAESKEIYR-----GYKMAEYHADIDYKSGDMKQDCDCEGGRIS-----H 81
Db 284 GQPAKKGKELVSPCLSPFSGEWEHAETVYRVSG--QKAAASLHELCAARVSEVLQNRVH 341
QY 82 QSDOKKIHY-----YGYMAYGPAQHAISTEK-----IKAKY 113
Db 342 RTEEVR-HVDFYAFSTYYDLAAGVG--LIDAEKGGSLVVGDFEIAKX 386

RESULT 6
US-09-608-285A-60
; Sequence 60, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
```

```
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-60

Query Match 10.5%; Score 70.5; DB 4; Length 471;
Best Local Similarity 29.6%; Pred. No. 1.4;
Matches 32; Conservative 12; Mismatches 31; Indels 33; Gaps 7;
QY 34 GAPAAESKEIYR-----GYKMAEYHADIDYKSGDMKQDCDCEGGRIS-----H 81
Db 312 GQPAKKGKELVSPCLSPFSGEWEHAETVYRVSG--QKAAASLHELCAARVSEVLQNRVH 369
QY 82 QSDOKKIHY-----YGYMAYGPAQHAISTEK-----IKAKY 113
Db 370 RTEEVR-HVDFYAFSTYYDLAAGVG--LIDAEKGGSLVVGDFEIAKX 414

RESULT 7
US-09-608-285A-27
; Sequence 27, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-27

Query Match 10.5%; Score 70.5; DB 4; Length 484;
Best Local Similarity 29.6%; Pred. No. 1.4;
Matches 32; Conservative 12; Mismatches 31; Indels 33; Gaps 7;
QY 34 GAPAAESKEIYR-----GYKMAEYHADIDYKSGDMKQDCDCEGGRIS-----H 81
```

Db 312 GQPAKCKEIVSPKSGEHEHAEVTVRVSG--QKAASLHEICARVSEVLQNRVH 369
QY 82 QSDCKEIVH-----YGYMAYGPAQHAISTEK-----IKAKY 113
Db 370 RTEEVK-HVDEYAFSYYYDLAAGVG--LIDAERKGSILVGDPEFIARVY 414

RESULT 8

US-09-370-265-27
; Sequence 27, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-27

Query Match 10.5%; Score 70.5; DB 4; Length 484;
Best Local Similarity 29.8%; Pred. No. 1.4;
Matches 32; Conservative 12; Mismatches 31; Indels 33; Gaps 7;
QY 34 GAPAESKEIVR-----GYKMAEYHADIVKVGSDMKOGDCDECLGGGRIS-----H 81
Db 312 GQPAKCKEIVSPKSGEHEHAEVTVRVSG--QKAASLHEICARVSEVLQNRVH 369
QY 82 QSDCKEIVH-----YGYMAYGPAQHAISTEK-----IKAKY 113
Db 370 RTEEVK-HVDEYAFSYYYDLAAGVG--LIDAERKGSILVGDPEFIARVY 414

RESULT 9

US-09-230-637-30
; Sequence 30, Application US/09230637
; Patent No. 6264958
; GENERAL INFORMATION:
; APPLICANT: Hayward, Gary
; APPLICANT: Nicholas, John
; APPLICANT: Hardwick, J. Marie
; APPLICANT: Reitz, Marvin
; TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma
; FILE REFERENCE: 1107.78372
; CURRENT APPLICATION NUMBER: US/09/230,637
; CURRENT FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 60/022,591
; PRIOR FILING DATE: 1996-07-25
; PRIOR APPLICATION NUMBER: PCT US 97/129311
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30

LENGTH: 333
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpes-like virus
US-09-230-637-30

Query Match 10.0%; Score 67.5; DB 4; Length 333;
Best Local Similarity 30.7%; Pred. No. 2;
Matches 31; Conservative 8; Mismatches 33; Indels 29; Gaps 6;

QY 15 DSDGVKTYLIRVHSAFSGAP-----AAESKEIVRGYKMA--ETHADIVKVGSDMKOG 68
Db 213 DQDG-----PADGAPVHRDSESVDEAGYKAGEPTNHDGRD--NVEPTAVG 259
QY 69 CDCCECLG-----GGRISHOSODKKIRHYGYSMAYGPA 100
Db 260 CDCNNLGAERRRATYCGGYVAGSDGAYSVSCINKA-GPS 299

RESULT 10

US-07-772-087-6
; Sequence 6, Application US/07772087
; Patent No. 5275945
; GENERAL INFORMATION:
; APPLICANT: HSIAO, Hung-Yu
; APPLICANT: FODGE, Douglas W.
; APPLICANT: LALONDE, James J.
; TITLE OF INVENTION: ALKALINE PROTEASES STABLE IN HEAVY-DUTY
; TITLE OF INVENTION: DETERGENT LIQUIDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: 10.07, Version #1.25
; APPLICATION NUMBER: US/07/772,087
; FILING DATE: 19911008
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16754/115 CHCO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacillus
; STRAIN: B. amyloliquefaciens
US-07-772-087-6

Query Match 9.7%; Score 65.5; DB 1; Length 381;
Best Local Similarity 22.8%; Pred. No. 4.5;
Matches 31; Conservative 15; Mismatches 53; Indels 37; Gaps 5;

QY 19 VKRYVILIRVHSAFSGAPAESKEIVRGYKMA--AEYHADIVKVGSDMKOG----- 67
Db 17 IFYMAFGSTSSAQAAGKSGEKKYIV-GFKQTMSTMSAAKRDVISEKGVQKQFKYVD 75
QY 68 -----GCDECLGGGRISHOSODKKIRHYGYSMAYG-----PAQH----- 102

Db 76 AASATLNKAVKELKDPSPVAAYVEDHVAHAAYASVPYGVSOIKAPALHSOGYTSNNKV 135
 QY 103 AISTEKIKAKPYDEV 118
 Db 136 AVIDSGIDSHDPLKV 151

RESULT 11

US-08-173-508-12
 ; Sequence 12, Application US/08173508
 ; Patent No. 5616485

GENERAL INFORMATION:

APPLICANT: Bartfeld, Daniel
 APPLICANT: Butler, Michael J.
 APPLICANT: Hadary, Dany
 APPLICANT: Jenish, David
 APPLICANT: Krieger, Timothy
 TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
 TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
 TITLE OF INVENTION: POLYPEPTIDES
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W.
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/173,508
 FILING DATE: 23-DEC-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 18740/125/CACO

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202 672 5300

TELEFAX: 202 672 5399

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 381 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-173-508-12

Query Match

Best Local Similarity 22.8%; Score 65.5; DB 1; Length 381;

Matches 31; Conservative 15; Mismatches 53; Indels 37; Gaps 5;

QY 19 VFVKYLIRVHAPSGAPAAESKEIVRGYK-----AEYHADIDYKVSQDMQO----- 67

Db 17 IFTMARFSTSSAQAAGSNGEKTYIV-GFKQTMSTMSAAKKDVISSEKGVOKQFYVD 75

QY 68 -----GDCCECLGGGRISHOSQDKIHYGYSMAYG-----PAQH----- 102

Db 76 AASATLNKAVKELKDPSPVAAYVEDHVAHAAYASVPYGVSOIKAPALHSOGYTSNNKV 135

QY 103 AISTEKIKAKPYDEV 118

Db 136 AVIDSGIDSHDPLKV 151

RESULT 12

US-08-265-310-12

; Sequence 12, Application US/08265310

; Patent No. 5856166

GENERAL INFORMATION:

APPLICANT: Bartfeld, Daniel
 APPLICANT: Butler, Michael J.
 APPLICANT: Hadary, Dany
 APPLICANT: Jenish, David
 APPLICANT: Krieger, Timothy
 APPLICANT: Malek, Lawrence T.
 APPLICANT: Soostmeyer, Gisela
 APPLICANT: Walczyk, Eva
 APPLICANT: Krysman, Phyllis
 APPLICANT: Garven, Sheila
 TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
 TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
 TITLE OF INVENTION: POLYPEPTIDES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W.
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/265,310
 FILING DATE: 24-JUN-1994
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/173,508

FILING DATE: 23-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 18740/133/CACO

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202 672 5300

TELEFAX: 202 672 5399

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 381 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-265-310-12

Query Match

Best Local Similarity 22.8%; Score 65.5; DB 2; Length 381;

Matches 31; Conservative 15; Mismatches 53; Indels 37; Gaps 5;

QY 19 VFVKYLIRVHAPSGAPAAESKEIVRGYK-----AEYHADIDYKVSQDMQO----- 67

Db 17 IFTMARFSTSSAQAAGSNGEKTYIV-GFKQTMSTMSAAKKDVISSEKGVOKQFYVD 75

QY 68 -----GDCCECLGGGRISHOSQDKIHYGYSMAYG-----PAQH----- 102

Db 76 AASATLNKAVKELKDPSPVAAYVEDHVAHAAYASVPYGVSOIKAPALHSOGYTSNNKV 135

QY 103 AISTEKIKAKPYDEV 118

Db 136 AVIDSGIDSHDPLKV 151

RESULT 13

US-08-951-742-12

; Sequence 12, Application US/08951742

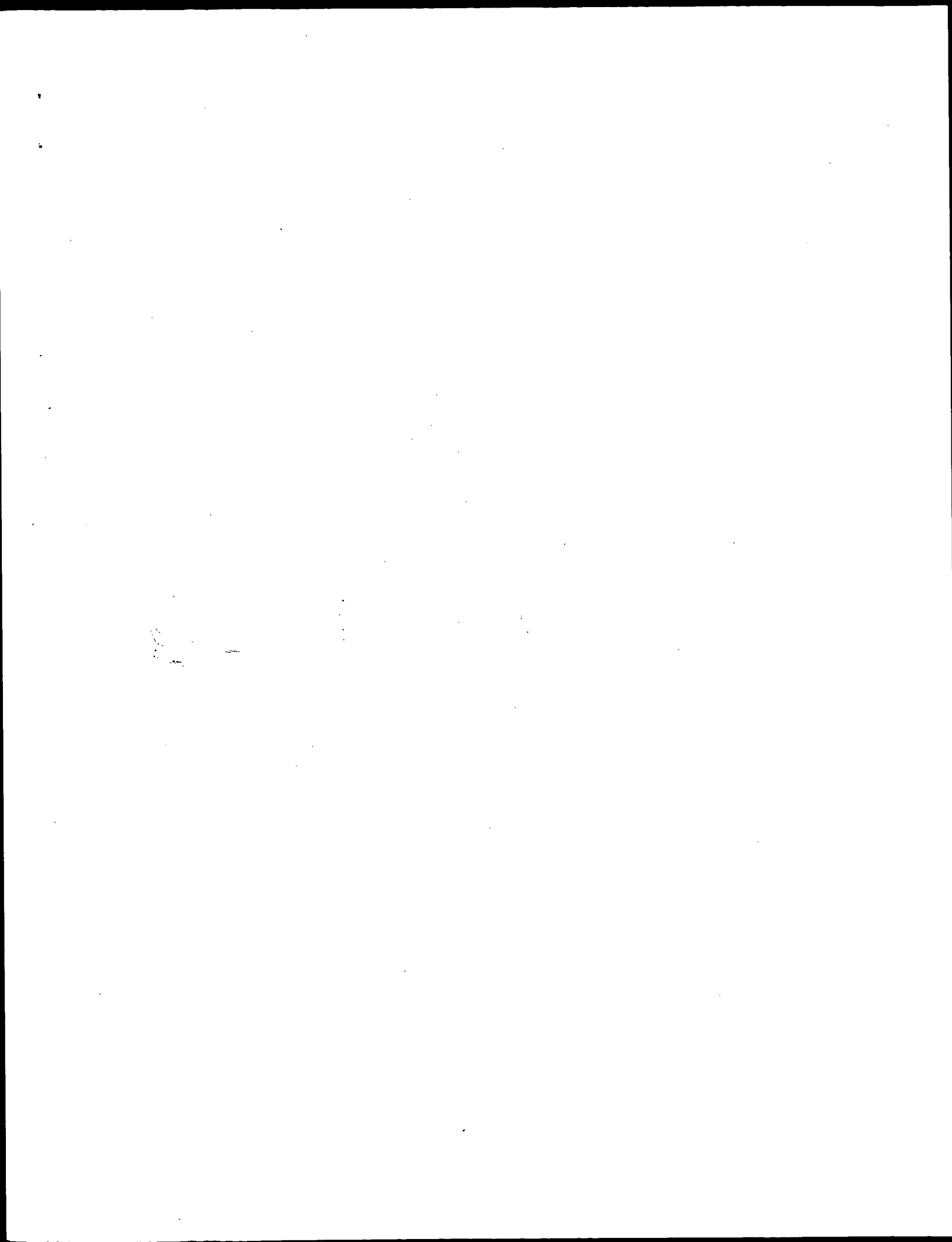
; Patent No. 6127144

; GENERAL INFORMATION:

; APPLICANT: Bartfeld, Daniel

; APPLICANT: Michael J. Butler

; APPLICANT: Dany Hadary



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:00:25 ; Search time 86.5826 Seconds

(Without alignments)
297.472 Million cell updates/sec

Title: US-09-914-831-2

Perfect score: 674

Sequence: 1 MAYADALALPVDVDSGVF.....TEKIKAKYPDYEVWANDGY 125

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP. Archaea:
2: SP. Bacteria:
3: SP. Fungi:
4: SP. Human:
5: SP. Invertebrate:
6: SP. Mammal:
7: SP. MHC:
8: SP. Organelle:
9: SP. Phage:
10: SP. Plant:
11: SP. Rodent:
12: SP. Virus:
13: SP. Vertebrate:
14: SP. Unclassified:
15: SP. Virus:
16: SP. Bacteriophage:
17: SP. Archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	674	100.0	125	4	Q9NRX4
2	663	98.4	125	4	Q9NRX4
3	563	83.5	124	11	Q9DAK9
4	248	36.8	135	5	Q9BH48
5	243	36.1	135	5	Q9BM97
6	243	36.1	135	5	Q9BM96
7	239	35.5	135	5	Q9BM95
8	238	35.3	119	5	Q9BM98
9	238	35.3	135	5	Q9BM99
10	237	35.2	135	5	Q9BM97
11	229	32.8	104	5	Q9BM97
12	219	32.5	109	5	Q9BM97
13	213	31.6	104	5	Q9BM97
14	211	31.3	102	5	Q9BM97
15	211	31.3	102	5	Q9BM97
16	193.5	28.7	115	5	P90861

17	169	25.1	188	4	Q9P019	Q9P019 homo sapien
18	162.5	24.1	160	5	Q9VIG0	Q9VIG0 drosophila
19	144	21.4	140	5	Q9BM90	Q9BM90 drosophila
20	142	21.1	140	5	Q9BM94	Q9BM94 drosophila
21	142	21.1	140	5	Q9BM91	Q9BM91 drosophila
22	140	20.8	140	5	Q9BM95	Q9BM95 drosophila
23	139	20.6	140	5	Q9BM93	Q9BM93 drosophila
24	139	20.6	140	5	Q9BM92	Q9BM92 drosophila
25	139	20.6	140	5	Q9BM82	Q9BM82 drosophila
26	137	20.3	140	5	Q9BM89	Q9BM89 drosophila
27	134	19.9	140	5	Q9BM85	Q9BM85 drosophila
28	134	19.9	140	5	Q9BM84	Q9BM84 drosophila
29	133	19.7	140	5	Q9BM83	Q9BM83 drosophila
30	131	19.4	140	5	Q9BM87	Q9BM87 drosophila
31	131	19.4	140	5	Q9BM86	Q9BM86 drosophila
32	131	19.4	140	5	Q9BM88	Q9BM88 drosophila
33	131	19.4	140	5	Q9BM88	Q9BM88 drosophila
34	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
35	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
36	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
37	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
38	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
39	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
40	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
41	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
42	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
43	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
44	125	19.4	140	5	Q9BM88	Q9BM88 drosophila
45	125	19.4	140	5	Q9BM88	Q9BM88 drosophila

ALIGNMENTS

RESULT 1	Q9NRX4	PRELIMINARY	PRT	125 AA.
ID	Q9NRX4	Q9NRX4		
AC	Q9NRX4	Q9NRX4		
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, last annotation update)			
DE	Sex-regulated protein janus-A (CGI-202) (Similar to HSPC141 protein).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=ADRENAL GLAND;			
RA	Li Y., Shi J., Huang C., Jiang C., Ren S., Zhou J., Yu Y., Xu S.,			
RA	Wang Y., Fu G., Chen Z., Han Z.;			
RT	* A novel gene expressed in human adrenal gland.;			
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=ADRENAL GLAND;			
RA	Peng Y., Gu Y., Li Y., Fu S., Gu J., Gu W., Jiang C., Yu Y., Han Z.,			
RA	Wang Y., Chen Z., Fu G.;			
RT	* A novel gene expressed in human adrenal gland.;			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=ADRENAL GLAND;			
RA	Strausberg R.;			
RA	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF164795; AAF80759.1; -			
DR	EMBL; AF285119; AAC01156.1; -			
DR	EMBL; BC024648; AAR24648.1; -			

SQ SEQUENCE 125 AA; 13832 MW; 24FC0CA2BADB78478 CRC64;
 Query Match 100.0%; Score 674; DB 4; Length 125;
 Best Local Similarity 100.0%; Pred. No. 2,1e-65;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAVADLALIPVDIDSDGVFKYVLIIRVHSAAPRSGAPAESEKEIVRGYKMAEYHADIDYDKV 60
 DB 1 MAVADLALIPVDIDSDGVFKYVLIIRVHSAAPRSGAPAESEKEIVRGYKMAEYHADIDYDKV 60
 QY 61 SGDMKQGCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDYEYTW 120
 DB 61 SGDMKQGCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDYEYTW 120
 QY 121 ANDGY 125
 DB 121 ANDGY 125
 RESULT 2
 Q9H0Y3 PRELIMINARY; PRT; 125 AA.
 AC Q9H0Y3; PRELIMINARY; PRT; 125 AA.
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 13.8 kDa protein.
 GN DKF2564M173.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansoorge W., Boescher M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Meves H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wandt R., Korn B., Klein M., Poustka A.;
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 DR EMBL: AL136644; CAB6579.1;
 KW Hypothetical protein.
 SQ SEQUENCE 125 AA; 13822 MW; 3C475F5B43ADE5A CRC64;
 Query Match 98.4%; Score 663; DB 4; Length 125;
 Best Local Similarity 97.6%; Pred. No. 3.2e-64;
 Matches 122; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAVADLALIPVDIDSDGVFKYVLIIRVHSAAPRSGAPAESEKEIVRGYKMAEYHADIDYDKV 60
 DB 1 MAVADLALIPVDIDSDGVFKYVLIIRVHSAAPRSGAPAESEKEIVRGYKMAEYHADIDYDKV 60
 QY 61 SGDMKQGCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDYEYTW 120
 DB 61 SGDMKQGCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDYEYTW 120
 QY 121 ANDGY 125
 DB 121 ANDGY 125
 RESULT 3
 Q9DAK9 PRELIMINARY; PRT; 124 AA.
 AC Q9DAK9; PRELIMINARY; PRT; 124 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE I700008C22Rik protein (RIKEN cDNA I700008C22 gene).

GN 1700008C22Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weller C., Whitaker C., Wilming L.,
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY GLAND;
 RX Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK005756; BAB24222.1;
 DR EMBL: BC028657; AAH28657.1;
 DR MGI: W1922704; I700008C22Rik.
 SQ SEQUENCE 124 AA; 13996 MW; 583F9A3CE95A2F66 CRC64;
 Query Match 83.5%; Score 563; DB 11; Length 124;
 Best Local Similarity 84.4%; Pred. No. 2.3e-53;
 Matches 103; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
 QY 4 ADLALIPVDIDSDGVFKYVLIIRVHSAAPRSGAPAESEKEIVRGYKMAEYHADIDYDKV 63
 DB 3 ADLALIPVDIDSDGVFKYVLIIRVHSAAPRSGAPAESEKEIVRGYKMAEYHADIDYDKV 62
 QY 64 MOKQGCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDYEYTW 123
 DB 63 LORNGYDCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDYEYTW 122
 QY 124 GY 125
 DB 123 GY 124
 RESULT 4
 Q9BH48 PRELIMINARY; PRT; 135 AA.
 AC Q9BH48; PRELIMINARY; PRT; 135 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janusa.
 GN JANA.
 OS Drosophila teissieri (Fruit fly), and
 OS Drosophila yakuba (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7243, 7245;
 RN [1]
 RP SEQUENCE FROM N.A.

RC SPECIES-D.teisleri, and D.yakuba;
RA Parsch J., Meiklejohn C.D., Hattl D.L.;
RT "Molecular evolution of the oculus and janus genes in the Drosophila
RT melanogaster species subgroup".
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY013342; AAG50363.1; -
DR EMBL: AY013341; AAG50362.1; -
DR FlyBase: FBgn0043602; DprikJana.
DR FlyBase: FBgn0043616; DdelJana.
SO SEQUENCE 135 AA: 1520 MW: C1756A84DE0571C CRC64;

Query Match	36.8%;	Score 248;	DB 5;	Length 135;
Best Local Similarity	42.4%;	Pred. No. 3.7e-19;		
Matches 53; Conservative	19;	Mismatches 47;	Indels	

```
QY 1 MAYVADLALIPDVYDSDGYFKYVLLRVHSAPRSGSPAASENSEIYRGKMMXEYHADIYDK 60
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 17 MSEAALAGVPLVHIPEGIFKYLTVN----IDGDA--SKAVYRGCDCCWADIIDRE 70
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 61 SCGMOKGCCCEELGGRISSHOSODKKIHYGYSMAYRPAQHAISTEKIKAPDYEVTW 120
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 71 EEFYFKTIGLAECPPGGRIEHNPEKYLKVIYGSQGFENADHQAOTKRILATKYPDYIET 130
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 121 ANDGY 125
Db 131 SDEGY 135
```

```
Query MatchScore      36.1%; Score 243; DB 5; Length 135;  
Best Local Similarity 40.8%; Pred. No. 1.3e-18;  
Matches    51; Conservative   20; Mismatches 48; Indels     6; Gaps    2;
```

DQ

1 MAVADLALIPVDIDSDGVFKYLIRVISAERSGAPAESKEIYRGKMAEYHADIYDKV 60
| : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
17 MSBEALAGPLVLIHSPEGIFRKYMTNV----IDGGDA--SKAVIRGRADCTWADTFDNR 70

DQ

61 SGDMOKQGCDCECLDGGRISHOSODKTIHYVGSMAYGPOMHALSTERIKAKYPDYEVTM 120
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
71 EEVFKKTLGIARACPGEGRIENPDKKYLKVGYSGQFGKDHAQTTRIILATKPDTITET 130

DQ

121 ANDGY 125
131 SDEGY 135

ID	09BM96	PRELIMINARY;	PRT;	135 AA.
AC	09BM96;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	JanusA.			
GN	JANuS.			
OS	Drosophila oreana (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Empidoidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7233;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21219192; PubMed=11319264;			
RA	Parsch J., MelkieJohn C.D., Hauschteck-Jungen E., Hunziker P.,			
RA	Hartl D.L.			
RT	"Molecular Evolution of the ocnus and janus Genes in the Drosophila			
RT	melanogaster Species Subgroup."			
RL	Mol. Biol. Evol. 18:801-811(2001).			
DR	EMBL; AY013344; AAC50365.1; -.			
DR	FLYbase; FBgn0043667; Dore\JanuA.			
SO	SEQUENCE 135 AA; 15146 MM; DB21E237BCD1307C CRC64;			

Query Match	36.1%;	Score 243;	DB 5;	Length 135;
Best Local Similarity	40.8%;	Pred. NO. 1.3e-18;		
Matches 51;	Conservative 20;	Mismatches 48;	Indels 6;	Gaps 2

QY	1	MAVADALLIPDVDDIDSDGYFKYVLLIRHNSAPRSGAPAAESKEIVRGKMAEYHADIDYKV	60
	17	MSEELALGVYLVHISEGIFPKYMINV----	1DGDGA--SKAVIRGADCTWADIDFDR
QY	61	SGDMOKGCGCDCELAGRISHOSODKKIHYGYSMAYGAPAOHAISTEKIKAKYDYEVTM	120
Db	71	EYVKKLIGLAECPGGGRIEHNPKKYLKYGYSGGFGKADHAKTITLTKIPDTIET	130
QY	121	ANDGY 125	
Db	131	SDEGY 135	
RESULT 7			
Q9BM98	ID	Q9BM98	PRELIMINARY; PRT; 135 AA.
AC	Q9BM98;		
DT	01-JUN-2001	(TREMBLrel. 17, Created)	
DT	01-JUN-2001	(Last sequence update)	
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)	
DE	JanuNA.		
GN	JANA.		
OS	Drosophila mauritiana (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;		
OC	Ephyroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7226;		
RP	SEQUENCE FROM N.A.		
FX	MEDLINE=21219192; PubMed=11319264;		
RA	Parisch J., Melkilejohn C.D., Hauschteck-Jungen E., Hunziker P.,		
RA	Harlt D.L.;		
RT	"Molecular Evolution of the <i>cenus</i> and <i>janus</i> Genes in the <i>Drosophila</i>		
RT	<i>melanogaster</i> Subgroup.";		
RL	Mol. Biol. Evol. 18:801-811(2001).		
EMBL	AY013340; AAC50361.1.		
DR	flyBase; Fbgn0043691; DmauJana.		
SO	SEQUENCE 135 AA; 15158 MW; 2F91E387BD70C1CC CRC64;		
Query Match			
	Best Local Similarity	35.5%; Score 239; DB 5; Length 135;	
	Matches 50; Conservative 21; Mismatches 48; Indels 6; Gaps 2;		
QY	1	MAVADALLIPDVDDIDSDGYFKYVLLIRHNSAPRSGAPAAESKEIVRGKMAEYHADIDYKV	60
	17	MSEELALGVYLVHISEGIFPKYMINV----	1DGDGA--SKAVIRGADCTWADIDFDR

RESULT 6
Q9BM96

```

Db 17 MSEBALAGVPLVHISPEGIFKYMINV----IDGDA--SKAVIRGFADCTWHDIFERE 70
Qy 61 SGMQKQGCDCCELGGRISHOSQDKKIHVYGSMAYPGPAHASTEKIKAPDYEVW 120
Db 71 EEVFKLGLRAECPPGGRIENPDKKYLKYGSGFGKADHQTIRLAKTIPDYIEI 130
Qy 121 ANDGY 125
Db 131 SDEGY 135

RESULT 8
087475 PRELIMINARY; PRT; 119 AA.
AC 087475;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DE AT12574p.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA Girarde R., Gonzalez M., Guatin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paclob J., Paragas V., Park S.,
RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY089320; AAL90058.1;
SQ SEQUENCE 119 AA; 13300 MW; 3FF85BCE91F8068 CRC64;

Query Match 35.3%; Score 238; DB 5; Length 119;
Best Local Similarity 40.0%; Pred. No. 3.9e-18;
Matches 50; Conservative 21; Mismatches 48; Indels 6; Gaps 2;

Qy 1 MAVADLALIPVDIDSDGVKRYLIRVHSA PRSGAPAAESKEIYRGKMAEYHADIDYKV 60
Db 1 MSEBALAGVPLVHISPEGIFKYMINV----IDGDA--SKAVIRGFADCTWHDIFERE 54
Qy 61 SGMQKQGCDCCELGGRISHOSQDKKIHVYGSMAYPGPAHASTEKIKAPDYEVW 120
Db 55 EEVFKLGLRAECPPGGRIENPDKKYLKYGSGFGKADHQTIRLAKTIPDYIEI 114
Qy 121 ANDGY 125
Db 115 SDEGY 119

RESULT 9
09BM99 PRELIMINARY; PRT; 135 AA.
AC 09BM99;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
GN JANUA.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RA Parsch J., Meikiejohn C.D., Hauschreck-Jungen E., Hunziker P.,
RA Hartl D.L.;
RL "Molecular Evolution of the ocnus and janus Genes in the Drosophila

```

```

RT melanogaster Species Subgroup.";
RL Mol. Biol. Evol. 18:801-811(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S36, S3, S4, S5, S6, S7, S8, S17, S19, S25, S31, AND S34;
RX MEDLINE=21518596; PubMed=11606541;
RA Parsch J., Meikiejohn C.D., Hartl D.L.;
RL "Patterns of DNA Sequence Variation Suggest the Recent Action of
RT Positive Selection in the janus-ocnus Region of Drosophila simulans.";
RN [1]
RP Genetics 159:647-657(2001).
RL EMBL: AY013339; AAG50360.1;
DR EMBL: AF393342; AAK72020.1;
DR EMBL: AF393331; AAK72009.1;
DR EMBL: AF393332; AAK72010.1;
DR EMBL: AF393333; AAK72011.1;
DR EMBL: AF393334; AAK72012.1;
DR EMBL: AF393335; AAK72013.1;
DR EMBL: AF393336; AAK72014.1;
DR EMBL: AF393337; AAK72015.1;
DR EMBL: AF393338; AAK72016.1;
DR EMBL: AF393339; AAK72017.1;
DR EMBL: AF393340; AAK72018.1;
DR EMBL: AF393341; AAK72019.1;
DR EMBL: FBgn0016348; Dsim\Jana.
SQ SEQUENCE 135 AA; 15206 MW; 2E2940304F7E21CD CRC64;

Query Match 35.3%; Score 238; DB 5; Length 135;
Best Local Similarity 40.0%; Pred. No. 4.5e-18;
Matches 50; Conservative 21; Mismatches 48; Indels 6; Gaps 2;

Qy 1 MAVADLALIPVDIDSDGVKRYLIRVHSA PRSGAPAAESKEIYRGKMAEYHADIDYKV 60
Db 17 MSEBALAGVPLVHISPEGIFKYMINV----IDGDA--SKAVIRGFADCTWHDIFERE 70
Qy 61 SGMQKQGCDCCELGGRISHOSQDKKIHVYGSMAYPGPAHASTEKIKAPDYEVW 120
Db 71 EEVFKLGLRAECPPGGRIENPDKKYLKYGSGFGKADHQTIRLAKTIPDYIEI 130
Qy 121 ANDGY 125
Db 131 SDEGY 135

RESULT 10
095YV7 PRELIMINARY; PRT; 135 AA.
AC 095YV7;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
GN JANUA.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RX MEDLINE=21518596; PubMed=11606541;
RA Parsch J., Meikiejohn C.D., Hartl D.L.;
RL "Patterns of DNA Sequence Variation Suggest the Recent Action of
RT Positive Selection in the janus-ocnus Region of Drosophila simulans.";
DR EMBL: AF393330; AAK72008.1;
DR EMBL: FBgn0016348; Dsim\Jana.
SQ SEQUENCE 135 AA; 15188 MW; 3F2940304E7E210A CRC64;

Query Match 35.2%; Score 237; DB 5; Length 135;
Best Local Similarity 40.8%; Pred. No. 5.8e-18;
Matches 49; Conservative 20; Mismatches 45; Indels 6; Gaps 2;

```


QY 6 LALIPVDIDSDGVFKYVILIRVHSAPRSGAPAESEKIVRGKMAEYHADIDYKSGDMQKGCDC 65
 DB 22 LAGVPLVHSPEGIFKTYMINV-----IDGDA--SKAVIRFADCTWHADIFEREVEVFK 75
 QY 66 KQGCDCBCLGGGRISHOSODKKIHVGYSMAYGPAQHAISTEKIKAKYDPYEVWANDGY 125
 DB 76 KLGRACPGGGRIRHNPKKYLKVGYSOGFGKADHAQTKRIATKYPDYTIETISDEGY 135

RESULT 11

Q9BMZ1 PRELIMINARY; PRT: 104 AA.
 ID 09BMZ1: PRELIMINARY; PRT: 104 AA.
 AC 09BMZ1: PRELIMINARY; PRT: 104 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila sechellia (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7238;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species."
 RL Genetics 156:1913-1931(2000).
 DR EMBL: AF284459; AAC49478.1; -;
 DR Flybase: FBgn0043637; DsecJana.
 FT NON_TER 1
 SQ SEQUENCE 104 AA; 11735 MW; 73BF32548B261427 CRC64;

Query Match 32.8%; Score 221; DB 5; Length 104;
 Best Local Similarity 40.4%; Pred. No. 2.3e-16;
 Matches 44; Conservative 19; Mismatches 40; Indels 6; Gaps 2;

QY 17 DGVFKYVILIRVHSAPRSGAPAESEKIVRGKMAEYHADIDYKSGDMQKGCDC 76
 DB 2 EGIFKTYMINV-----IDGDA--SKAVIRFADCTWHADIFEREVEVFK 75
 QY 77 GRISHOSODKKIHVGYSMAYGPAQHAISTEKIKAKYDPYEVWANDGY 125
 DB 56 GRLEHNPDKKYLKVGYSOGFGKADHAQTKRIATKYPDYTIETISDEGY 104

RESULT 12

Q9BMZ7 PRELIMINARY; PRT: 109 AA.
 ID 09BMZ7: PRELIMINARY; PRT: 109 AA.
 AC 09BMZ7: PRELIMINARY; PRT: 109 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7240;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species."
 RL Genetics 156:1913-1931(2000).

DR EMBL: AF284453; AAC49466.1; -;
 DR Flybase: FBgn0016348; DsimJana.
 FT NON_TER 1
 SQ SEQUENCE 109 AA; 12384 MW; 1118CE3722E56A6F CRC64;

Query Match 32.5%; Score 219; DB 5; Length 109;
 Best Local Similarity 39.5%; Pred. No. 4e-16;
 Matches 45; Conservative 19; Mismatches 44; Indels 6; Gaps 2;

QY 12 VDIDSDGVFKYVILIRVHSAPRSGAPAESEKIVRGKMAEYHADIDYKSGDMQKGCDC 71
 DB 2 WHISPEGIFKTYMINV-----IDGDA--SKAVIRFADCTWHADIFEREVEVFK 75
 QY 72 ECLGGGRISHOSODKKIHVGYSMAYGPAQHAISTEKIKAKYDPYEVWANDGY 125
 DB 56 ECPGGGRIRHNPKKYLKVGYSOGFGKADHAQTKRIATKYPDYTIETISDEGY 109

RESULT 13

Q9BMZ5 PRELIMINARY; PRT: 104 AA.
 ID 09BMZ5: PRELIMINARY; PRT: 104 AA.
 AC 09BMZ5: PRELIMINARY; PRT: 104 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7226;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species."
 RL Genetics 156:1913-1931(2000).
 DR EMBL: AF284456; AAC49472.1; -;
 DR Flybase: FBgn0043691; DmauJana.
 FT NON_TER 1
 SQ SEQUENCE 104 AA; 11834 MW; 690E994E3A961437 CRC64;

Query Match 31.6%; Score 213; DB 5; Length 104;
 Best Local Similarity 39.4%; Pred. No. 1.7e-15;
 Matches 43; Conservative 19; Mismatches 41; Indels 6; Gaps 2;

QY 17 DGVFKYVILIRVHSAPRSGAPAESEKIVRGKMAEYHADIDYKSGDMQKGCDC 76
 DB 2 EGIFKTYMINV-----IDGDA--SKAVIRFADCTWHADIFEREVEVFK 75
 QY 77 GRISHOSODKKIHVGYSMAYGPAQHAISTEKIKAKYDPYEVWANDGY 125
 DB 56 GRLEHNPDKKYLKVGYSOGFGKADHAQTKRIATKYPDYTIETISDEGY 104

RESULT 14

Q9BH68 PRELIMINARY; PRT: 102 AA.
 ID 09BH68: PRELIMINARY; PRT: 102 AA.
 AC 09BH68: PRELIMINARY; PRT: 102 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7240;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KENYA_5; AND KENYA_12;
 RX MEDLINE=20556156; PubMed-11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 Drosophila simulans Complex Species."
 RL Genetics 156:1913-1931(2000).
 DR EMBL; AF284454; AAG49468.1; -
 DR EMBL; AF284455; AAG49470.1; -
 DR Flybase; FBgn0016348; Dsim\Jana.
 FT NON_TER
 SQ SEQUENCE 102 AA; 11608 MW; 7C1AC80B7E9369EB CRC64;

Query Match 31.3%; Score 211; DB 5; Length 102;
 Best Local Similarity 39.8%; Pred. No. 2.7e-15;
 Matches 43; Conservative 18; Mismatches 41; Indels 6; Gaps 2;

QY 18 GVFKEYLIRVHSA PRSGAPAESEKEIVRGYKMAEYHADIDYKVS GDMOKGDCDCCECLGGG 77
 DB 1 GIFRYWMINV----IDGGDA--SKAVIRPADCTWADIFEREBEVFKLGLRAECPPGG 54
 QY 78 RISHOSODKKIHYGYSMAYGPAOHAISTEKIKAKYPDYEVTWANDGY 125
 DB 55 RIEHNPDKKYLKYGYSGGFGKADHAQTKRIATKYPDYTIETISDEGY 102

RESULT 15
 Q9BH67 PRELIMINARY; PRT; 102 AA.
 AC Q9BH67;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_taxid=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CAP MALHEUREUX; AND PORT-LOUIS;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 Drosophila simulans Complex Species."
 RL Genetics 156:1913-1931(2000).
 DR EMBL; AF284457; AAG49474.1; -
 DR EMBL; AF284458; AAG49476.1; -
 DR Flybase; FBgn0043691; Dmau\Jana.
 FT NON_TER
 SQ SEQUENCE 102 AA; 11608 MW; 7C1AC80B7E9369EB CRC64;

Query Match 31.3%; Score 211; DB 5; Length 102;
 Best Local Similarity 39.8%; Pred. No. 2.7e-15;
 Matches 43; Conservative 18; Mismatches 41; Indels 6; Gaps 2;

QY 18 GVFKEYLIRVHSA PRSGAPAESEKEIVRGYKMAEYHADIDYKVS GDMOKGDCDCCECLGGG 77
 DB 1 GIFYWMINV----IDGGDA--SKAVIRPADCTWADIFEREBEVFKLGLRAECPPGG 54
 QY 78 RISHOSODKKIHYGYSMAYGPAOHAISTEKIKAKYPDYEVTWANDGY 125
 DB 55 RIEHNPDKKYLKYGYSGGFGKADHAQTKRIATKYPDYTIETISDEGY 102

Search completed: May 7, 2003, 19:06:35
 Job time : 88.5826 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 18:54:40 ; Search time 14.9083 Seconds

(without alignments)
347.763 Million cell updates/sec

Title: US-09-914-831-2

Perfect score: 674
Sequence: 1 MAVADLALIPVDIDSDGVF.....TEKIRAKYPDYEVWANDGY 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276	40.9	146	1	JANA_DROPS
2	242	35.9	135	2	JANA_DROME
3	153	22.7	140	1	JANB_DROPS
4	142	21.1	140	1	JANB_DROME
5	83	12.3	444	1	TBB_EUPFO
6	81	12.0	444	1	TBB2_DAUCA
7	80	11.9	423	1	TBB2_PORPU
8	80	11.9	444	1	TBB_ACHKL
9	80	11.9	444	1	TBB_PHYCI
10	80	11.9	451	1	TBB_NABGR
11	80	11.9	467	1	TBB1_PHYPO
12	79.5	11.8	186	1	DHML_METEX
13	79	11.7	441	1	TBB_BABBO
14	79	11.7	444	1	TBB1_HUMAN
15	79	11.7	444	1	TBB7_CHICK
16	79	11.7	444	1	TBBX_HUMAN
17	79	11.7	444	1	TBB_EUPOC
18	79	11.7	445	1	TBB1_SOYBN
19	79	11.7	445	1	TBB_PLAFA
20	79	11.7	446	1	TBB1_TRIVI
21	79	11.7	446	1	TBB_EUPCR
22	79	11.7	451	1	TBB4_PORPU
23	78	11.6	341	1	TBB_HALDI
24	78	11.6	386	1	TBB1_AVEBA
25	78	11.6	408	1	TBB3_SOYBN
26	78	11.6	411	1	TBB2_ANEPH
27	78	11.6	440	1	TBB3_PEA
28	78	11.6	442	1	TBB1_PARTE
29	78	11.6	442	1	TBB_STYLE
30	78	11.6	443	1	TBB2_XENLA
31	78	11.6	443	1	TBB_TETPY
32	78	11.6	443	1	TBB_TETTH
33	78	11.6	444	1	TBB1_ORYSA

34	78	11.6	444	1	TBB2_MAIZE	P18026 zea mays (m
35	78	11.6	444	1	TBB9_ARATH	P29517 arabidopsis
36	78	11.6	445	1	TBB1_CHICK	P09203 gallus gall
37	78	11.6	445	1	TBB1_ELEIN	G02900 eleusine in
38	78	11.6	445	1	TBB1_GADMO	G03963 gadus morhu
39	78	11.6	445	1	TBB1_RAT	P04691 rattus norv
40	78	11.6	445	1	TBB1_WHEAT	G02902 triticum ae
41	78	11.6	445	1	TBB2_CHICK	P32882 gallus gall
42	78	11.6	445	1	TBB2_HUMAN	P05217 homo sapien
43	78	11.6	445	1	TBB3_CHICK	P09206 gallus gall
44	78	11.6	445	1	TBB3_MAIZE	Q43695 zea mays (m
45	78	11.6	445	1	TBB3_WHEAT	G02900 triticum ae

ALIGNMENTS

RESULT 1	JANA_DROPS	STANDARD:	PRT:	146 AA.
AC	P54364:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Sex-regulated protein janus-A.			
GN	JANA.			
OS	Drosophila pseudoobscura (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7237;			
RM	[1]			
RX	SEQUENCE FROM N.A.			
RP	MEDLINE=95214615; PubMed=7700229;			
RA	Yanickoatas C., Ferrer P., Vincent A., Lepesant J.-A.;			
RT	*Separate cis-regulatory sequences control expression of serendipity			
RT	beta and janus A, two immediately adjacent Drosophila genes.*;			
RL	Mol. Gen. Genet. 246:549-560(1995).			
CC	-1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-B.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL: S77099; AAB3912.2; -			
DR	FLYBase: FBgn0015151; DpseJana.			
DR	SEQUENCE 146 AA; 16291 MW; 02F7DEC570070D7B CRC64;			
SO	SEQUENCE			
Query Match	40.9%; Score 276; DB 1; Length 146;			
Best Local Similarity	45.7%; Pred. No. 1.3e-22;			
Matches	59; Conservative 20; Mismatches 38; Indels 12; Gaps 3;			
OY	1 MAVADLALIPVDIDSDGVFVKYLIRVHSAPRSGAPAAE-----SKETIRGKMAEYVADI 56			
DB	26 MSDQDLALIPVDIDSDGVFVKYLIRV-----TKETRADGTEPSKLVYRGYADC---ADI 77			
OY	57 YDKVSGDMQKGCDCGGRISHQSDKKIHYVGYSMAYGPAQHAISTEKIRAKYPDY 116			
DB	78 YERDGTGKGLDTEFCGGGRTEHNEPKRYLKYVGHSTGYGKADHAEKRVLLTKYKNY 137			
OY	117 EYTWANDGY 125			
DB	138 EETSDEGT 146			
RESULT 2				
JANA_DROME	STANDARD:	PRT:	135 AA.	
ID	P20346; Q9VAB6;			

RESULT 4
JANB_DROME STANDARD: PRT: 140 AA.
ID JANB_DROME
AC P20349: 09VAB7;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sex-regulated protein janus-B.
GN JANB OR CG7931.
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-Canton-S.
RX MEDLINE=89343970; PubMed=2503707;
RA Yancoskas C., Vincent A., Lepesant J.-A.;
RT "Transcriptional and posttranscriptional regulation contributes to
the sex-regulated expression of two sequence-related genes at the
janus locus of Drosophila melanogaster.";
RL Mol. Cell. Biol. 9:2526-2535(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abdel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Bertram J.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherly J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos K., Delcher A.L., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden Klamos I., Simpson M.C., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: GERM LINE CELLS OF ADULT MALES.
CC -1- DEVELOPMENTAL STAGE: FROM THE THIRD-INSTAR LARVAL STAGE TO THE
CC ADULT STAGE.
CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-A.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M27033; AAC34204.1; -
DR EMBL: AE003772; AAF56996.1; -
DR PIR: B32317; B32317.
DR Flybase: FBgn0001281; janB.
SQ SEQUENCE 140 AA; 15859 MW; 5750CC28BH61D7FE CRC64;
Query Match 21.1%; Score 142; DB 1; Length 140;
Best Local Similarity 29.2%; Pred. No. 2,9e-08;
Matches 33; Conservative 30; Mismatches 42; Indels 8; Gaps 4;
OY 9 IPDVDDISDGVKRYVIRVHSAFPGAPAAESKEIVRGYKMAVHADIVKYSQDMOROG 68
DB 31 VPRVKI-TKGONRYLLVINH---THGFTKRGYGVIRGAD-VDNHLAVPDSILBELPEEG 84
OY 69 COCECGGGGRISHOSQDKIHVGYSMVGPQQAISTEKIA--KYDYEVT 119
DB 85 ICAKILGGGRILNEAFNRKIKIYGTSTFGADHRTNRLQAWTYYRDKFIT 137
RESULT 5
TBB_EUPRO STANDARD: PRT: 444 AA.
ID TBB_EUPRO
AC 09N2M6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tubulin beta chain (Beta-tubulin).
OS Euplotis focardii.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
OC Euplotida; Euplotidae; Euplotes.
ON NCBI_TaxID=36767;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TNI;
RX MEDLINE=94372955; PubMed=8087111;
RA Miceli C., Ballarín P., Di Giuseppe G., Valbonesi A., Luporini P.;
RT "Identification of the tubulin gene family and sequence determination
of one beta-tubulin gene in a cold-poikilotherm protozoan, the
RT antarctic ciliate Euplotes focardii.";
RL J. Eukaryot. Microbiol. 41:420-427(1994).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S72098; AAB31932.1; -
DR InterPro: IPR002453; beta_tubulin.
DR InterPro: IPR000158; FtsZ.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; tubulin_FtsZ.
DR Pfam: PF00091; tubulin.1.
DR PRINTS: PR00423; CELDVISFTSZ.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
KW Microtubules; GTP-binding.
FT NP_BIND 140 145 GTP (POTENTIAL).

SQ SEQUENCE 444 AA; 49807 MW; 1F50DD2FA60B9DB06 CRC64;
 Query Match 12.3%; Score 83; DB 1; Length 444;
 Best Local Similarity 28.4%; Pred. No. 0.22;
 Matches 21; Conservative 11; Mismatches 28; Indels 14; Gaps 2;
 QY 46 GYKMAEYH----ADIDYKVSQDMQKQDCGCGGRIHQSDKKIHVYGYSMAYGPAQ 101
 DB 98 GNNWAKGHTEGAEILDSVLDVYKAEKNECCLQGFOYTH-----SLGGGTSGSM 147
 QY 102 HAISTEKIKAKYPD 115
 DB 148 GTLLISKIREKYPD 161
 RESULT 6
 TBB2_DAUCA STANDARD; PRT; 444 AA.
 AC 039697;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Tubulin beta-2 chain (Beta-2 tubulin).
 GN TUBB2.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Kintoki;
 RX MEDLINE=97338494; PubMed=9195054;
 RA Okamura S., Naito K., Somehara S., Ohkawa H., Kuramori S., Tatsuta M.,
 RA Minamizono M., Kataoka T.;
 RT Characterization of the carrot beta-tubulin gene coding a divergent
 RT isotype, beta-2.";
 RL Cell Struct. Funct. 22:291-298(1997).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- TISSUE SPECIFICITY: FOUND IN AREAS OF RAPIDLY DIVIDING TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U63927; AAB64308.1;
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin.1.
 DR PRINTS: PRO1161; TUBULIN.1.
 DR PROSITE: PS00227; TUBULIN.1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG.1.
 KM Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 140 146 GTP (POTENTIAL).
 SQ SEQUENCE 444 AA; 49571 MW; 359D3C7433DB188 CRC64;
 Query Match 12.0%; Score 81; DB 1; Length 444;
 Best Local Similarity 27.0%; Pred. No. 0.37;
 Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;
 QY 46 GYKMAEYH----ADIDYKVSQDMQKQDCGCGGRIHQSDKKIHVYGYSMAYGPAQ 101
 DB 98 GNNWAKGHTEGAEILDSVLDVYKAEKNECCLQGFOYCH-----SLGGGTSGSM 147

QY 102 HAISTEKIKAKYPD 115
 DB 148 GTLLISKIREKYPD 161
 RESULT 7
 TBB2_PORPU STANDARD; PRT; 423 AA.
 AC P50260;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE Tubulin beta-2 chain (Fragment).
 GN TUBB2.
 OS Porphyra purpurea.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
 OX NCBI_TaxID=2787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Avonport;
 RA Mackay R.M., Gallant J.W.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 267992; CA91940.1;
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin.1.
 DR PROSITE: PS00227; TUBULIN.1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG. PARTIAL.
 KM Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 115 121 GTP (POTENTIAL).
 SQ SEQUENCE 423 AA; 47367 MW; 4EB030756DD3A45A CRC64;
 Query Match 11.9%; Score 80; DB 1; Length 423;
 Best Local Similarity 27.0%; Pred. No. 0.44;
 Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;
 QY 46 GYKMAEYH----ADIDYKVSQDMQKQDCGCGGRIHQSDKKIHVYGYSMAYGPAQ 101
 DB 73 GNNWAKGHTEGAEILDSVLDVYKAEKNECCLQGFOYTH-----SLGGGTSGSM 122
 QY 102 HAISTEKIKAKYPD 115
 DB 123 GTLLISKIREKYPD 136
 RESULT 8
 TBB2_ACHKL STANDARD; PRT; 444 AA.
 AC P20802;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DE Tubulin beta chain.
 OS Achlya klebsiana.
 OC Eukaryota; stramenopiles; Oomycetes; Saprolegniales; Saprolegniaceae;
 OC Achlya.
 OX NCBI_TaxID=4767;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368709; PubMed=2394720;
RA Cameron L.E., Huttsul J.-A., Thorlacius L., Lejohm H.B.;
RT "Cloning and analysis of beta-tubulin gene from a protoctist.";
RL J. Biol. Chem. 265:15245-15252(1990).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DEVELOPMENTAL STAGE: SPOROPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05597; AAA63161.1; -.
DR PIR: A35885; A35885.
DR InterPro: IPR002453; Beta.tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin_1.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN_1.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG_1.
DR Microtubules: GTP-binding.
DR NP_BIND 139 145 GTP (POTENTIAL).
SQ SEQUENCE 444 AA; 49851 MW; 7B3BF25B29626B5E CRC64;

Query Match 11.9%; Score 80; DB 1; Length 444;
Best Local Similarity 27.0%; Pred. No. 0.47;
Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;

QY 46 GYKAEYH---ADYDKVSGMOKGCCGCGGGRISHQSDKKIHVYGYSMAYGPAQ 101
DB 97 GNNWAKGHTEGAELIDSVLDVYKKAESCDLQGFQITH-----SLGGGTGSGM 146
QY 102 HAISTEKIRAKYPD 115
DB 147 GTLLISKIREXPD 160

RESULT 9
TBL_PHYCI
AC 059837; STANDARD; PRT; 444 AA.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta chain.
OS Phytophthora cinamoml.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
NCBI_TaxID=4785;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DAR 52646;
RA Weerakoon N.D., Roberts J.K., Lehnen L.P. Jr., Wilkinson J.M.,
RA Marshall J.S., Hardham A.R.;
RT "Isolation and characterization of the single beta-tubulin gene in
RT Phytophthora cinamoml.";
RL Mycologia 90:85-95(1998).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U22050; AAC05441.1; -.
DR InterPro: IPR002453; Beta.tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin_1.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN_1.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG_1.
DR Microtubules: GTP-binding.
DR NP_BIND 140 146 GTP (POTENTIAL).
SQ SEQUENCE 444 AA; 49922 MW; 8392C8C288378929 CRC64;

Query Match 11.9%; Score 80; DB 1; Length 444;
Best Local Similarity 27.0%; Pred. No. 0.47;
Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;

QY 46 GYKAEYH---ADYDKVSGMOKGCCGCGGGRISHQSDKKIHVYGYSMAYGPAQ 101
DB 98 GNNWAKGHTEGAELIDSVLDVYKKAESCDLQGFQITH-----SLGGGTGSGM 147
QY 102 HAISTEKIRAKYPD 115
DB 148 GTLLISKIREXPD 161

RESULT 10
TBL_NAEGR
AC P34108; STANDARD; PRT; 451 AA.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1994 (Rel. 38, Last annotation update)
DE Tubulin beta chain.
OS Naegleria gruberi.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OC Naegleria gruberi.
NCBI_TaxID=5762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEG;
RX MEDLINE=94118284; PubMed=8289261;
RA Lal E.Y., Remillard S.P., Fulton C.;
RT "A beta-tubulin gene of Naegleria encodes a carboxy-terminal
RT tyrosine. Aromatic amino acids are conserved at carboxy termini.";
RL J. Mol. Biol. 235:377-388(1994).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z13961; CAAT8362.1; -.
DR PIR: S30514; S30514.
DR InterPro: IPR002453; Beta.tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.

```

DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PR01161; TUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 KW Microtubules: GTP-binding.
 FT NP_BIND 140 146 GTP (POTENTIAL).
 SQ SEQUENCE 451 AA; 50511 MW; 842777FC706BDD1 CRC64;
 Query Match 11.9%; Score 80; DB 1; Length 451;
 Best Local Similarity 27.0%; Pred. No. 0.48;
 Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;
 QY 46 GYKMAEYH---ADYDKSGDMOGCCCELCGGGRISHOSQDKKHVGYGSMAYGPAQ 101
 DB 98 GNNMAGHYTEGALIDVLDVVRKEAESCDCLOGFOIAH-----SLGGTSGSM 147
 QY 102 HAISTEKIKAKYPD 115
 DB 148 GTLLISKIREYPD 161
 RESULT 11
 TBHL_PHYPO STANDARD; PRT; 467 AA.
 AC P07436;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tubulin beta-1 chain.
 GN BETA AND BETB.
 OS Physarum polycephalum (Slime mold).
 OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
 OC Physarum.
 OX NCBI_TaxID=5791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J0352;
 RA Paul E.C.A., Buchsachner G.L. Jr., Cunningham D.B., Dove W.F.,
 Burland T.G.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-445 FROM N.A.
 RC STRAIN=J0352;
 RA MEDLINE=92211323; PubMed=1556551;
 RA Paul E.C.A., Buchsachner G.L. Jr., Cunningham D.B., Dove W.F.,
 Burland T.G.;
 RT "Preferential expression of one beta-tubulin gene during flagellate
 development in Physarum."
 RL J. Gen. Microbiol. 138:229-238(1992).
 RN [3]
 RP SEQUENCE OF 4-207 FROM N.A.
 RC STRAIN=CL;
 RA MEDLINE=88271316; PubMed=3391166;
 RA Wierskiold A.K., Poetsch B., Haugli F.;
 RT "Cloning and expression of a beta tubulin gene of Physarum
 polycephalum."
 RL Eur. J. Biochem. 174:491-495(1988).
 RN [4]
 RP PARTIAL SEQUENCE OF 1-217; 234-262 AND 277-286.
 RC MEDLINE=87080317; PubMed=3539596;
 RA Singhofer-Wowra M., Clayton L., Dawson P., Gull K., Little M.;
 RT "Amino-acid sequence data of beta-tubulin from Physarum polycephalum
 myxamebae."
 RL Eur. J. Biochem. 161:669-679(1986).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SUBCELLULAR LOCATION: MITOSIS IN THE SLIME MOLD PLASMODIUM DIFFERS
 FROM THE PROCESS IN MANY EUKARYOTES. THE TUBULIN CHAINS MUST BE
 TRANSPORTED TO THE NUCLEI FOR INTRANUCLEAR ASSEMBLY OF THE
 SPINDLE.
 CC -1- DEVELOPMENTAL STAGE: BETA IS PREFERENTIALLY EXPRESSED IN

CC FLAGELLATE AND BETA IN AMEBA.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: M58521; AAA29974.1; -
 DR EMBL: X12371; CAA30932.1; -
 DR PIR: A44848; A44848.
 DR PIR: A25342; A25342.
 DR PIR: S02532; S02532.
 DR InterPro: IPR002453; Beta.tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PR01161; TUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 KW Microtubules; GTP-binding; Multigene family; Nuclear protein.
 FT NP_BIND 140 146 GTP (POTENTIAL).
 FT VARIANT 39 39 E -> D (IN BETB).
 FT VARIANT 283 283 S -> A (IN BETB).
 FT CONFLICT 165 165 C -> D (IN REF. 4).
 FT CONFLICT 196 196 A -> T (IN REF. 4).
 FT CONFLICT 238 238 C -> S (IN REF. 4).
 SQ SEQUENCE 467 AA; 52134 MW; BA2C330A6FDC964 CRC64;
 Query Match 11.9%; Score 80; DB 1; Length 467;
 Best Local Similarity 27.0%; Pred. No. 0.5;
 Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;
 QY 46 GYKMAEYH---ADYDKSGDMOGCCCELCGGGRISHOSQDKKHVGYGSMAYGPAQ 101
 DB 98 GNNMAGHYTEGALIDVLDVVRKEAESCDCLOGFOIAH-----SLGGTSGSM 147
 QY 102 HAISTEKIKAKYPD 115
 DB 148 GTLLISKIREYPD 161
 RESULT 12
 DHML_METEX STANDARD; PRT; 186 AA.
 AC P00372; O60146;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Methylamine dehydrogenase light chain precursor (EC 1.4.99.3) (MADH).
 GN MADH.
 OS Methylbacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylbacterium group; Methylbacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1 / NCIMB 9133;
 RA MEDLINE=91358185; PubMed=1653226;
 RA Chistoserdov A.Y., Tsygankov Y.D., Lidstrom M.E.;
 RT "Genetic organization of methylamine utilization genes from
 Methylbacterium extorquens AM1."
 RL J. Bacteriol. 173:5901-5908(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1 / NCIMB 9133;
 RA MEDLINE=94292425; PubMed=8021187;
 RA Chistoserdov A.Y., Chistoserdova L.V., McIntire W.S., Lidstrom M.E.;
 RT "Genetic organization of the mau gene cluster in Methylbacterium
 extorquens AM1: complete nucleotide sequence and generation and

RT Characteristics of mau mutants."
 RL J. Bacteriol. 176:4052-4065(1994).
 [3]
 RP SEQUENCE OF 58-186 FROM N.A.
 RC STRAIN-AM1 / NCIMB 9133;
 RA MEDLINE=91025043; PubMed=2121141;
 RX Chistoserdo A.Y., Tsyanov Y.D., Lidstrom M.E.;
 RT "Cloning and sequencing of the structural gene for the small subunit
 of methyamine dehydrogenase from Methylobacterium extorquens AM1:
 evidence for two tryptophan residues involved in the active center."
 RL Biochem. Biophys. Res. Commun. 172:211-216(1990).
 [4]
 RP SEQUENCE OF 58-186.
 RC STRAIN-AM1 / NCIMB 9133;
 RA MEDLINE=83186062; PubMed=6841324;
 RX Ishii Y., Hase T., Fukumori Y., Matsubara H., Tobari J.;
 RT "Amino acid sequence studies of the light subunit of methyamine
 dehydrogenase from Pseudomonas AM1: existence of two residues binding
 the prosthetic group."
 RL J. Biochem. 93:107-119(1983).
 [5]
 RP CHARACTERIZATION OF COFACTOR.
 RX MEDLINE=91227905; PubMed=2028257;
 RA McIntire W.S., Wenner D.E., Chistoserdo A.Y., Lidstrom M.E.;
 RT "A new cofactor in a prokaryotic enzyme: tryptophan tryptophylquinone
 as the redox prosthetic group in methyamine dehydrogenase."
 RL Science 252:817-824(1991).
 CC -1- FUNCTION: METHYLAMINE DEHYDROGENASE CARRIES OUT THE OXIDATION OF
 METHYLAMINE. ELECTRONS ARE PASSED FROM METHYLAMINE DEHYDROGENASE
 TO AMICYANIN.
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + acceptor = RCHO + NH(3)
 + reduced acceptor
 CC -1- COFACTOR: EACH OF THE SMALL SUBUNITS CONTAINS A COVALENTLY BOUND
 POQ-LIKE COFACTOR CALLED TRYPTOPHAN TRYPTOPHYLQUINONE (TTO).
 CC -1- PATHWAY: Methyamine utilization.
 CC -1- SUBUNIT: TETRAMER OF TWO LIGHT AND TWO HEAVY CHAINS.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: HIGH, TO OTHER SPECIES MACH LIGHT CHAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M57963; AAA68894.1; -
 DR EMBL; L26406; AAB46936.1; -
 DR EMBL; M58517; AAA25379.1; -
 DR PIR; A36676; DEPSNL.
 DR HSP; P22619; 2BBK.
 DR InterPro: IPR004229; Me-amine-deh_L.
 DR Pfam; PF02975; Me-amine-deh_L; 1.
 KW Oxidoreductase; Electron transport; Periplasmic; TTO; Signal.
 FT SIGNAL 1 57
 FT CHAIN 58 186 METHYLAMINE DEHYDROGENASE LIGHT CHAIN.
 FT BINDING 112 112 TTO.
 FT BINDING 163 163
 FT DISULFID 78 143 BY SIMILARITY.
 FT DISULFID 84 116 BY SIMILARITY.
 FT DISULFID 91 176 BY SIMILARITY.
 FT DISULFID 93 141 BY SIMILARITY.
 FT DISULFID 101 132 BY SIMILARITY.
 FT DISULFID 133 164 BY SIMILARITY.
 FT VARIANT 106 106 K -> L.
 FT CONFLICT 74 74 D -> N (IN REF. 4).
 FT SEQUENCE 186 AA; 20084 MW; 4F8504FEFF0331D9 CRC64;
 Query Match 11.8%; Score 79.5; DB 1; Length 186;
 Best Local Similarity 25.5%; Pred. No. 0.2;
 Matches 39; Conservative 14; Mismatches 43; Indels 57; Gaps 8;

QY 2 AVADLAIIPVDVDSGVFKYVLLIRHNSADPSGAPAESEIVRGYKMAEYHADIDKVS 61
 DB 34 AVAGVALVPLIPVDRRG-----RV-----SRNAAEASGDRG-KMKPQNDV----- 75
 QY 62 GDMKQKCCD-----CECLGGGRIS-----HOSDKKIHYGVS 94
 DB 76 -----QSCDYWRHNCISIDGNICDSCGSLTSCPPGTKLASSSWVASCYNPTDKOSYLISVR 130
 QY 95 MAYPAQH-----AISTEKIRAKYPDEYTWAND 123
 DB 131 DCCGANYSGCACLNTE---GELPYRPERGND 160
 RESULT 13
 TBB_BABBO STANDARD; PRT; 441 AA.
 ID TBB_BABBO
 AC 004709;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Tubulin beta chain.
 OS Babesia bovis.
 OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.
 OX NCBI_TaxID=5865;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93341585; PubMed=8341332;
 RA Casu R.E.;
 RT "Sequence of a cDNA encoding beta-tubulin from Babesia bovis."
 RL Mol. Biochem. Parasitol. 59:339-340(1993).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L00978; AAA27796.1; -
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam; PF00091; tubul.n; 1.
 DR PRINTS; PRO1161; TUBULIN.
 DR PROSITE; PS00227; TUBULIN; 1.
 DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
 KW Microtubules; GTP-binding.
 FT NP BIND 140 146 GMP (BY SIMILARITY).
 FT SEQUENCE 441 AA; 49392 MW; 412DIE7CFE3030E CRC64;
 Query Match 11.7%; Score 79; DB 1; Length 441;
 Best Local Similarity 27.0%; Pred. No. 0.59;
 Matches 20; Conservative 12; Mismatches 28; Indels 14; Gaps 2;
 QY 46 GYKMAEYH---ADIDYKSGDMKQKCCCECLGGGRISHOSQDKKIHYGYSMAYPAQ 101
 DB 98 GNNWAGHTEGAEILDVLDVVRKEAEGCDLGGFOITH-----SLGGGTGSGM 147
 QY 102 HAISTEKIRAKYPD 115
 DB 148 GTLLISKIRKREYPD 161
 RESULT 14
 TBB1_HUMAN STANDARD; PRT; 444 AA.
 ID TBB1_HUMAN
 AC P07437;

```

DN 01-Apr-1988 (Rel. 07, Created)
DT 01-Apr-1988 (Rel. 07, Last sequence update)
DT 15-Jun-2002 (Rel. 41, Last annotation update)
DE Tubulin beta-1 chain.
DE TUBBL1.
GN TUBBL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=63232883; Pubmed=6688039;
RA Hall J.L., Dudley L., Dobner P.R., Lewis S.A., Cowan N.J.;
RT "Identification of two human beta-tubulin isotypes.";
RL Mol. Cell. Biol. 3:854-862(1983).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- DOMAIN: THE HIGHLY ACIDIC CARBOXYL-TERMINAL REGION MAY BIND
CC CATIONS SUCH AS CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00314; AAB59507.1; -.
DR PIR: A26561; A26561.
DR Genew; HGNC:12412; TUBB.
DR MIM; 191130; -.
DR InterPro; IPR002453; Beta_tubulin.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_FtsZ.
DR Pfam; PF00091; tubulin: 1.
DR PRINTS; P01161; TUBULIN.
DR PROSITE; PS00227; TUBULIN_1.
DR PROSITE; PS00228; TUBULIN_B_Autoreg; 1.
DR Microtubules; GTP-binding; Multigene family.
DR NP_BIND 140 146 GTP (POTENTIAL).
DR TUBBL1
DR SEQUENCE 444 AA; 49759 MW; B941CE8BA40DA9D CRC64;
SQ
Query Match 11.7%; Score 79; DB 1; Length 444;
Best Local Similarity 22.0%; Pred. NO. 0.6;
Matches 31; Conservativity 18; Mismatches 36; Indels 56; Gaps 6;
OY 11 DVDIDSDGFE-----KVV-----LIVHSAP-----RSGAPAA 38
DB 41 DIQIDRISVYVTEAGKGVPRALLIVDLFEGTMDSVRSRPFQIIRPDNFVVGSGA--- 97
OY 39 ESKETVIGKMAEYH---ADIVYKVSQDMQKQDCDECLGSGGRISHSQDKKIHYGYS 94
DB 98 -----GNMNAKGHYTEAEALVDSVLVDVVRKEAESCDLQAGFOLTH-----SLG 140
OY 95 MAYGPAQHAISTEKIKAKYPD 115
DB 141 GGTGSGMGTLISLKIREPYD 161

```

```
AC P09244; DT 01-MAR-1989 (Rel. 10, Created) DT 01-MAR-1989 (Rel. 10, Last sequence update) DT 15-JUL-1999 (Rel. 38, Last annotation update) DE Tubulin beta-7 chain ('Tubulin beta 4'). OS Gallus gallus (Chicken). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; CC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; CC Gallus. OX NCBI_Taxid=9031; RN [1] RP SEQUENCE FROM N.A. RX MEDLINE=68172491; PubMed=3351937; RA Montelino M.J., Cleveland D.W.; RT "Sequence of chicken c beta 7 tubulin. Analysis of a complete set of vertebrate beta-tubulin isoforms." J. Mol. Biol. 199; 439:446(1988). RL -I- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGABLE SITE ON THE BETA CHAIN AND ONE AT A NONEXCHANGABLE SITE ON THE ALPHA-CHAIN. CC -I- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS. CC -I- SIMILARITY: BELONGS TO THE TUBULIN FAMILY. ----- THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb.sib.ch). ----- CC EMBL: X07011; CAA30060.1; -. DR PIR: S01713; S01713. DR InterPro: IPR002453; Beta_tubulin. DR InterPro: IPR000217; Tubulin. DR InterPro: IPR003008; Tubulin_FtsZ. DR Pfam: PF00091; tubulin; 1. DR PRINTS: PR01161; TUBULIN. DR PROSITE: PS00227; TUBULIN_1. DR PROSITE: PS00228; TUBULIN_B AUTOREG: 1. KW Microtubules; GTP-binding; Multigene family. FT NP_BIND 140 146 GTP (POTENTIAL). SO SEQUENCE 444 AA; 49671 MW; 1BCDDA36773A103 CRC64;
```

```
OY Query Match          11.7%; Score 79; DB 1; Length 444;  
Db Best Local Similarity 22.0%; Pred. No. 0.6;  
Matches 31; Conservative 18; Mismatches 36; Indels 56; Gaps 6,  
  
OY      11 DAVIDSCGF-----KYV-----LIYNHAP-----RSGNAPA 38  
    | :| :||  
DB      41 DLQADRISVTNGATGYPPRAIIIVDEPGTMDSVRGGPGQGIFRPDNIYFGNSGA--- 97  
  
OY      39 ESKEITVGKMAEYH----ADIYDKVSQMGOKGCDCGLCGISHOSODKKIHVYGYS 94  
    ||:||: |:| :||:||||:  
DB      98 ----GNWNAIGHTEGAEVLVDSDLVYVKREASCCCLGPFQLTH-----SLG 140  
  
OY      95 MAYGPQAQHAISTEKIKAYPD 115  
    | :| :||  
DB     141 GGTSGMGTLLISKIREEPD 161
```

Search completed: May 7, 2003, 19:03:56
Job time : 15.9083 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:00 ; Search time 26.9495 Seconds

(without alignments)
445,900 Million cell updates/sec

Title: US-09-914-831-2

Perfect score: 674

Sequence: 1 MAVADLALIPVDIDSDGVF.....TEKIRAKIPDEVTWANDGY 125

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	265.5	39.4	149	2	S53876
2	242	35.9	135	2	A32317
3	193.5	28.7	115	2	T21833
4	142	21.1	140	2	B32317
5	141	20.9	140	2	S53877
6	80	11.9	204	2	S02532
7	80	11.9	282	2	A25342
8	80	11.9	444	2	A35885
9	80	11.9	445	2	A44848
10	80	11.9	451	2	S30514
11	79.5	11.8	186	1	DEPSNL
12	79	11.7	437	2	I38369
13	79	11.7	444	2	S01713
14	79	11.7	444	2	S18456
15	79	11.7	444	2	A26561
16	79	11.7	444	2	S31400
17	79	11.7	444	2	E25437
18	79	11.7	445	2	A44949
19	79	11.7	445	2	JAC048
20	79	11.7	446	2	B30309
21	78	11.6	383	2	S42480
22	78	11.6	386	2	S14570
23	78	11.6	411	2	S32669
24	78	11.6	434	2	S52008
25	78	11.6	440	2	S20870
26	78	11.6	441	2	S45071
27	78	11.6	442	2	S25182
28	78	11.6	442	2	S00683
29	78	11.6	443	2	S05966

30	78	11.6	443	2	S01769	tubulin beta-2 cha
31	78	11.6	443	2	S01768	tubulin beta-1 cha
32	78	11.6	443	2	S41470	tubulin beta chain
33	78	11.6	444	2	S35191	tubulin beta-1 cha
34	78	11.6	444	2	J01593	tubulin beta-9 cha
35	78	11.6	444	2	S14702	tubulin beta-2 cha
36	78	11.6	444	2	JC2518	tubulin beta chain
37	78	11.6	444	2	S18457	beta-tubulin R2242
38	78	11.6	444	2	S18457	tubulin beta chain
39	78	11.6	445	1	UBPCB	tubulin beta chain
40	78	11.6	445	1	UBPCB	tubulin beta chain
41	78	11.6	445	1	UBPCB	tubulin beta chain
42	78	11.6	445	2	T08726	tubulin beta chain
43	78	11.6	445	2	J00120	tubulin beta chain
44	78	11.6	445	2	A25113	tubulin beta chain
45	78	11.6	445	2	A24701	tubulin beta-3 cha

ALIGNMENTS

RESULT 1
S53876
sex-regulated protein janus A - fruit fly (Drosophila pseudobscura)
C:Species: Drosophila pseudobscura
C:Date: 14-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Jan-1999
C:Accession: S53876
R:Janicostas, C.; Ferrer, P.; Vincent, A.; Lepesant, J.A.
Mol. Gen. Genet. 246, 549-560, 1995
A:Title: Separate cis-regulatory sequences control expression of serendipity beta and
A:Reference number: S53876; MUID:95214615; PMID:7700229
A:Accession: S53876
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <YAN>
A:Cross-references: GB:S77099; NID:g913977; PID:g913978
C:Genetics:
A:Gene: janA
A:Cross-references: FlyBase:FBgn0015151
A:Introns: 78/1; 119/3

Query Match 39.4%; Score 265.5; DB 2; Length 149;
Best Local Similarity 45.0%; Pred. No. 2.9e-20;
Matches 58; Conservative 17; Mismatches 45; Indels 9; Gaps 2;

QY 1 MAVADLALIPVDIDSDGVFVKYILRVHSAPRSQAPAE-----SKETVRCYKMAEYADI 56
DB 26 MSDQALAKIPVDIDSDGVFVKYILRV-----TGKTADGTGFSKLVGVADCGWADI 80
QY 57 YDKVSGDMQKQDCDCGCGGGRISHOSODKKIHVYGYSMAYGPAOHAISTEKIRAKYPDY 116
DB 81 YGRGTGKTIGTGLDNCGLGGRGRIHNPGRKYLKAYGHSTGTCKADHAGSKVLLTKTKNY 140
QY 117 EYTWANDGY 125
DB 141 GIGTSDGY 149

RESULT 2
A32317
sex-regulated protein janA - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 29-Jan-1990 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C:Accession: A32317
R:Janicostas, C.; Vincent, A.; Lepesant, J.A.
Mol. Cell. Biol. 9, 2526-2535, 1989
A:Title: Transcriptional and posttranscriptional regulation contributes to the sex-re
A:Reference number: A32317; MUID:89343970; PMID:2503707
A:Accession: A32317
A:Status: preliminary
A:Molecule type: DNA; mRNA
A:Residues: 1-135 <YAN>
A:Cross-references: GB:M27033; NID:g3522914; PIDN:AMC34203.1; PID:g387597

A: Note: the authors translated the codon TNG for residue 98 as Met
 C: Genetics:
 A: Gene: FlyBase: JanB
 A: Cross-references: FlyBase: FBgn0001280

Query Match 35.9%; Score 242; DB 2; Length 135;
 Best Local Similarity 40.0%; Pred. No. 7.2e-18;
 Matches 50; Conservative 22; Mismatches 47; Indels 6; Gaps 2;

QY 1 MAVADIALPVDIDSDGVFRKYLIRVHSAFNSGAPPAESKEIVRGYKMAEYHADIDYKV 60
 DB 17 MSEALAGVGLVHISPEGIKRYMINVF---DGDA--SKAVIRGFADCTWADIFERE 70
 QY 61 SGMOKGCGCEGLGGRISHSQDKKIHYGSMAYGPAQNAISTEKIKAKYPDYETW 120
 DB 71 EEFVKRLGLAECPPGGRIHNPCKIKLYKYSQGFGRADHAKOTRIKATKPDYITIEI 130
 QY 121 ANDGY 125
 DB 131 SDEGY 135

RESULT 3
 T21833
 hypothetical protein F36A2.8 - *Caenorhabditis elegans*

C: Species: *Caenorhabditis elegans*
 C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C: Accession: T21833
 R: Lennard, N.
 submitted to the EMBL Data Library, October 1996

A: Reference number: Z19476
 A: Accession: T21833
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1115 <MIL>
 A: Cross-references: EMBL: Z21077; PTDN: CAB03070.1; GSPDB: GN00019; CESP: F36A2.8
 A: Experimental source: clone F36A2
 C: Genetics:
 A: Gene: CESP: F36A2.8
 A: Map position: 1
 A: Introns: 22/3; 52/3

Query Match 28.7%; Score 193.5; DB 2; Length 115;
 Best Local Similarity 42.5%; Pred. No. 6.6e-13;
 Matches 37; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 39 ESKRIYRGYKMAEYHADIDYKVSQDKGDCGCGRISHSQDKRIHYGYSMAYG 98
 DB 30 QOKFVVRGYYRCSFHDILQETKSSAPD-LKLKCVGGRIKIDELGKIDILVYGYSTGYG 88
 QY 99 PAQHAISTEKIKAKYPDYETWANDGY 125
 DB 89 RADHQIADVILKOKYRPYNIHFNSNDGY 115

RESULT 4
 B32317
 sex-regulated protein JanB - fruit fly (*Drosophila melanogaster*)
 C: Species: *Drosophila melanogaster*
 C: Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 21-Jul-2000
 C: Accession: B32317
 R: Yanicostas, C.; Vincent, A.; Lepesant, J.A.
 Mol. Cell. Biol. 9, 2526-2535, 1989
 A: Title: Transcriptional and posttranscriptional regulation contributes to the sex-regul
 A: Reference number: A32317; MUID: 89343970; PMID: 2503707

A: Accession: B32317
 A: Status: preliminary
 A: Molecule type: DNA; mRNA
 A: Residues: 1-140 <YAN>
 A: Cross-references: GB: M27033; NID: g3522914; PTDN: AAC34204.1; PID: g387598
 C: Genetics:
 A: Gene: FlyBase: JanB
 A: Cross-references: FlyBase: FBgn0001281

Query Match 21.1%; Score 142; DB 2; Length 140;
 Best Local Similarity 29.2%; Pred. No. 1.8e-07;
 Matches 33; Conservative 30; Mismatches 42; Indels 8; Gaps 4;

QY 9 IPVDVIDSDGVFRKYLIRVHSAFNSGAPPAESKEIVRGYKMAEYHADIDYKVSQDKG 68
 DB 31 VPRVKI-TKGQRKLVVNIH---THGFTKGRVYRGAD-VDNHLAVFDSILELPEEG 84
 QY 69 CMCCEGLGGRISHSQDKKIHYGSMAYGPAQNAISTEKIKAKYPDYETW 119
 DB 85 ICATILGGGRILNEAKNKIKIYGTSTRTFGADHTFRNIIQAWTYTKDEKIT 137

RESULT 5
 S53877
 sex-regulated protein Janus B - fruit fly (*Drosophila pseudoobscura*)

C: Species: *Drosophila pseudoobscura*
 C: Date: 18-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Jan-1999
 C: Accession: S53877
 R: Yanicostas, C.; Ferrer, P.; Vincent, A.; Lepesant, J.A.
 Mol. Gen. Genet. 246, 549-560, 1995
 A: Title: Separate cis-regulatory sequences control expression of serendipity beta and
 A: Reference number: S53876; MUID: 95214615; PMID: 7700229
 A: Accession: S53877
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-140 <YAN>
 A: Cross-references: GB: S77099; NID: g913977; PID: g913979
 C: Genetics:
 A: Gene: JanB
 A: Cross-references: FlyBase: FBgn0014784
 A: Introns: 18/1; 70/1; 111/3

Query Match 20.9%; Score 141; DB 2; Length 140;
 Best Local Similarity 32.2%; Pred. No. 2.3e-07;
 Matches 39; Conservative 23; Mismatches 45; Indels 14; Gaps 6;

QY 4 ADLALIPVDIDSDGVFRKYLIRV--HSAPNSGAPPAESKEIVRGYKMAEYHADIDYKVS 61
 DB 26 ADLKLKPRVDI-KGCKRLRLILSYIIRGTR-----HARTVVRGMN-TDSHDIDYKKNV 77
 QY 62 GDMOKGDCGCEGLGGRISHSQDKKIHYGSMAYGPAQNAISTEKI---KATYPDYEV 118
 DB 78 RAMGRKIGLCTKICGCKMDNSARKIGIHGCKTFGAANH-KTKGILLSSSRYKFNFI 136

QY 119 F 119
 DB 137 F 137

RESULT 6
 S02532
 tubulin beta-1 chain - slime mold (*Physarum polycephalum*) (fragment)
 C: Species: *Physarum polycephalum*
 C: Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 23-Feb-1997
 C: Accession: S02532
 R: Werenskiold, A.K.; Poetsch, B.; Haugli, F.
 Eur. J. Biochem. 174, 491-495, 1988
 A: Title: Cloning and expression of a beta tubulin gene of *Physarum polycephalum*.
 A: Reference number: S02532; MUID: 88271316; PMID: 3391166
 A: Accession: S02532
 A: Molecule type: DNA
 A: Residues: 1-204 <WER>
 A: Cross-references: EMBL: X12371
 C: Genetics:
 A: Gene: beta1
 A: Introns: 95/1
 C: Superfamily: tubulin
 C: Keywords: nucleus

Query Match 11.9%; Score 80; DB 2; Length 204;
 Best Local Similarity 27.0%; Pred. No. 0.76;

Db 148 GTLLISKIREYD 161

RESULT 11

DEPSNT

amine dehydrogenase (EC 1.4.99.3) light chain - Methylolacterium extorquens

M:Alternate names: methylamine dehydrogenase light chain

C:Species: Methylolacterium extorquens

C:Date: 14-Nov-1983 #sequence_revision 17-Nov-1995 #text_change 11-Jun-1999

C:Accession: A36676; A00384; B38123

R:Chitoseirov, A.Y.; Tsygankov, Y.D.; Lidstrom, M.E.

Biochem. Biophys. Res. Commun. 172, 211-216, 1990

A:Title: Cloning and sequencing of the structural gene for the small subunit of methylamine center.

A:Reference number: A36676; MUID:91025043; PMID:2121141

A:Accession: A36676

A:Molecule type: DNA

A:Residues: 58-186 <CH3>

A:Cross-references: GB:M58517; NID:9150012; PIDN:AAA25379.1; PID:9150013; GB:M38387

R:Shill, Y.; Hase, T.; Fukumori, Y.; Matsubara, H.; Tobari, J.

J. Biochem. 93, 107-119, 1983

A:Title: Amino acid sequence studies of the light subunit of methylamine dehydrogenase.

A:Reference number: A00384; MUID:83186062; PMID:6843324

A:Accession: A00384

A:Molecule type: protein

A:Residues: 58-73, 'N', 75-105, 'L', 107-111, 'X', 113-162, 'X', 164-186 <IS2>

A:Experimental source: AM1

A:Note: 106-Lys was also found

R:Chitoseirov, A.Y.; Lidstrom, M.E.

J. Bacteriol. 173, 5909-5913, 1991

A:Title: The small-subunit polypeptide of methylamine dehydrogenase from Methylolacterium

A:Reference number: A38123; MUID:91358386; PMID:1885555

A:Accession: B38123

A:Molecule type: DNA

A:Residues: 1-59 <CH2>

A:Cross-references: GB:M57963

C:Comment: The active enzyme is a tetramer of two light and two heavy chains.

C:Superfamily: amine dehydrogenase light chain

C:Keywords: oxidoreductase; quinoprotein

F:112-163/Cross-link: tryptophan-tryptophyl quinone (Trp-Trp) #status predicted

F:112/Modified site: tryptophyl quinone (Trp) #status predicted

Query Match 11.8%; Score 79.5; DB 1; Length 186;

Best Local Similarity 25.5%; Pred. No. 0.79;

Matches 39; Conservative 14; Mismatches 43; Indels 57; Gaps 8;

Db 2 AVADLALIPVDIDSDGVFKYVLRVHSAPRGAPAAESKEIKYKMAEYADIDKYS 61

Db 34 AVAGVALVPLLPVDRRG-----RV-----SRANAESAAGDPRG-KMKRQDNV----- 75

QY 62 GDMQKQSCD-----CECLGGGRIS-----HQSODRKIHVYGS 94

Db 76 -----QSCDYMWRHCSIDGNICDGGSLTSCPGTKLASSWVASCYNPDKOSTYLISIR 130

QY 95 MAYGPAQH-----AISTEKIRAKYDYEWAND 123

Db 131 DCCGANYSGRACALNTE---GELPYVRPERGND 160

RESULT 12

beta-tubulin - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 13-Aug-1999

C:Accession: I38369

R:Hall, J.L.; Dudgeon, L.; Dobner, P.R.; Lewis, S.A.; Cowan, N.J.

Mol. Cell. Biol. 3, 854-862, 1983

A:Title: Identification of two human beta-tubulin isoforms.

A:Reference number: I38369; MUID:83244582; PMID:6865944

A:Accession: I38369

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-437 <RES>
A:Cross-references: EMBL:V00599; NID:937490; PIDN:CAA33844.1; PID:9897763
C:Superfamily: tubulin

Query Match

Best Local Similarity 11.7%; Score 79; DB 2; Length 437;

Matches 31; Conservative 18; Mismatches 36; Indels 56; Gaps 6;

QY 11 DVDIDSDGVF-----KYV-----LIRVHSAP-----RSGAPAA 38

Db 34 DLQDRISVYVNEATGKRYPRAILVLEPGTMDVSRSRPGQIFRPDNEVFGQSGA--- 90

QY 39 ESKELVRYKMAEYH---ADYDKVSGDMQKQGCDCCLGGRISHQSODRKIHVYGS 94

Db 91 -----GNNAKGHYTGAEIVDSVLDVYKKEAESDCDCLQGFLTH-----SLG 133

QY 95 MAYGPAQHAIISTEKIRAKYPD 115

Db 134 GGTGSGMGTLLISKIREYD 154

RESULT 13

tubulin beta-7 chain - chicken

C:Species: Gallus gallus (chicken)

C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 13-Aug-1999

C:Accession: S01713

R:Monteiro, M.J.; Cleveland, D.W.

J. Mol. Biol. 199, 439-446, 1988

A:Title: Sequence of chicken c-beta-7 tubulin. Analysis of a complete set of vertebrate

A:Reference number: S01713; MUID:88172491; PMID:3351937

A:Accession: S01713

A:Molecule type: mRNA

A:Residues: 1-444 <MON>

A:Cross-references: EMBL:X07011; NID:963166; PIDN:CAA30060.1; PID:963167

C:Superfamily: tubulin

Query Match 11.7%; Score 79; DB 2; Length 444;

Best Local Similarity 22.0%; Pred. No. 2.3;

Matches 31; Conservative 18; Mismatches 36; Indels 56; Gaps 6;

QY 11 DVDIDSDGVF-----KYV-----LIRVHSAP-----RSGAPAA 38

Db 41 DLQDRISVYVNEATGKRYPRAILVLEPGTMDVSRSRPGQIFRPDNEVFGQSGA--- 97

QY 39 ESKELVRYKMAEYH---ADYDKVSGDMQKQGCDCCLGGRISHQSODRKIHVYGS 94

Db 98 -----GNNAKGHYTGAEIVDSVLDVYKKEAESDCDCLQGFLTH-----SLG 140

QY 95 MAYGPAQHAIISTEKIRAKYPD 115

Db 141 GGTGSGMGTLLISKIREYD 161

RESULT 14

tubulin beta chain (clone 16T) - Chinese hamster

C:Species: Crictetus griseus (Chinese hamster)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999

C:Accession: S18456

R:Ahmad, S.; Singh, B.; Gupta, R.S.

Biochim. Biophys. Acta 1090, 252-254, 1991

A:Title: Nucleotide sequences of three different isoforms of beta-tubulin cDNA from C

A:Reference number: S18456; MUID:92031702; PMID:1657186

A:Accession: S18456

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-444 <AHN>

A:Cross-references: EMBL:X60784; NID:949480; PIDN:CAA3197.1; PID:949481

C:Superfamily: tubulin

Query Match 11.7%; Score 79; DB 2; Length 444;

Best Local Similarity 22.0%; Pred. No. 2.3;

Matches 31; Conservative 18; Mismatches 36; Indels 56; Gaps 6;

```

QY 11 DVDIDSQGVF-----KYV-----LIRVHSAP-----RSGAPAA 38
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 41 DLQDLRISVYVNEKRTGKYVPRAILVLDLEPQTMDSVRSRPFQIFRPDNEVFQSGA--- 97
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 39 ESKETIVRGYKMAEYH---ADIDKVSQDMQKQCCDCECLGGGRISHOSODKKIHVYGY 94
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 98 -----GNNMAKGHYTEGAEVLDVSVLVYVYRKEAESCDCLQGFOLTH-----SLG 140
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 95 MAYGPAQHAISTEKIKAKYPD 115
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 141 GGTSQGMGTLISKIREYPD 161
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 15

A26561

tubulin beta chain - human

C:Species: Homo sapiens (man)

C:Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 13-Aug-1999

C:Accession: A26561

R:Lee, M.G.S.; Lewis, S.A.; Wilde, C.D.; Cowan, N.J.

Cell 33, 477-487, 1983

A:Reference number: A26561; MUID:83232883; PMID:6688039

A:Accession: A26561

A:Molecule type: mRNA

A:Residues: 1-444 <LE>

A:Cross-references: GB:000314; NID:q338694; PIDN:AAB59507.1; PID:q338695

A>Note: the authors translated the codon GAG for residue 111 as Gly

C:Superfamily: tubulin

Query Match 11.7%; Score 79; DB 2; Length 444;
 Best Local Similarity 22.0%; Pred. No. 2.3;

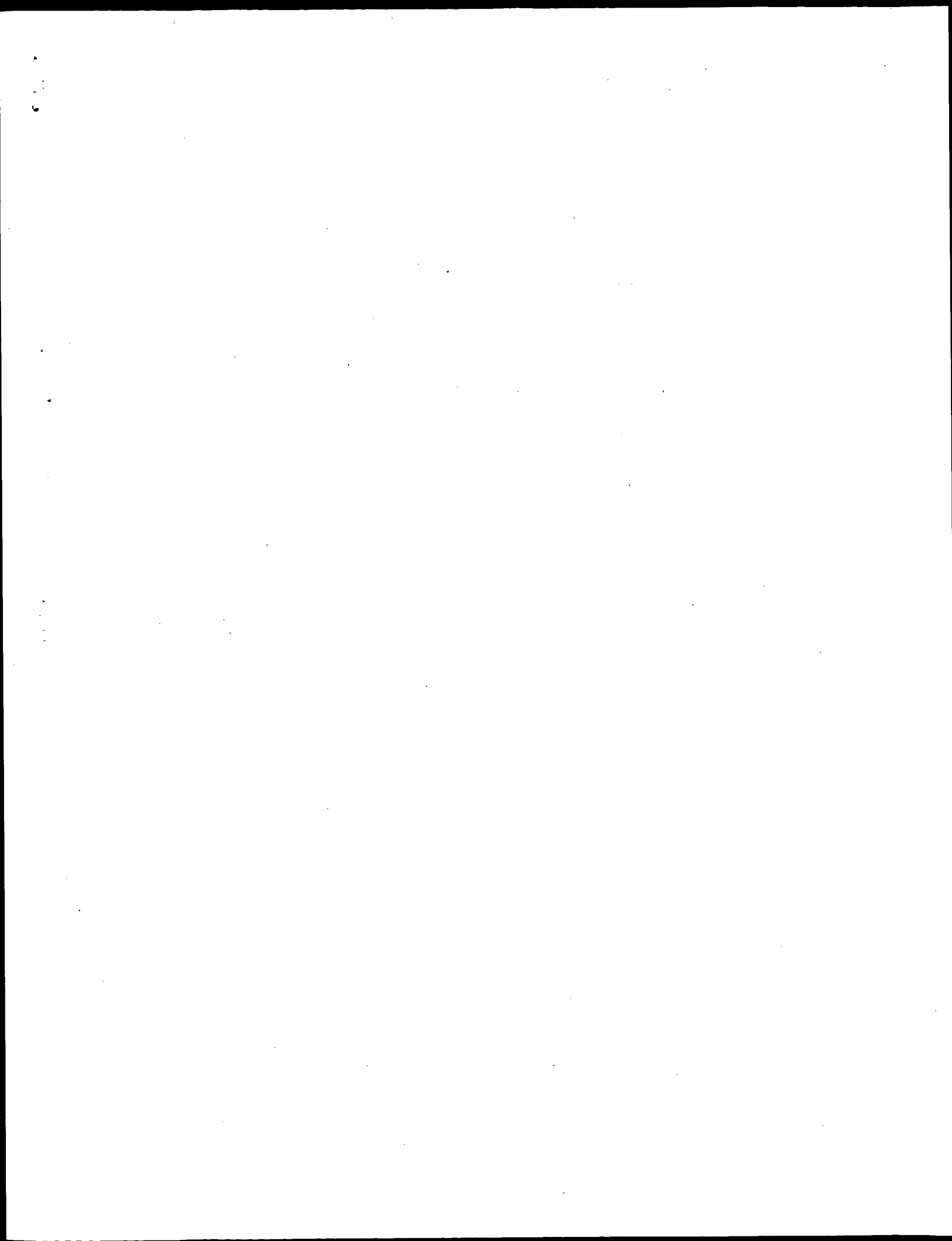
Matches 31; Conservative 18; Mismatches 36; Indels 56; Gaps 6;

```

QY 11 DVDIDSQGVF-----KYV-----LIRVHSAP-----RSGAPAA 38
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 41 DLQDLRISVYVNEKRTGKYVPRAILVLDLEPQTMDSVRSRPFQIFRPDNEVFQSGA--- 97
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 39 ESKETIVRGYKMAEYH---ADIDKVSQDMQKQCCDCECLGGGRISHOSODKKIHVYGY 94
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 98 -----GNNMAKGHYTEGAEVLDVSVLVYVYRKEAESCDCLQGFOLTH-----SLG 140
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 95 MAYGPAQHAISTEKIKAKYPD 115
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 141 GGTSQGMGTLISKIREYPD 161
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Search completed: May 7, 2003, 19:07:29
 Job time : 28.9495 secs



PT	-
XX	
PS	Claim 5; Page 17; 39pp; English

CC The present sequence represents human histidine phosphatase polypeptide
CC
CC The polypeptide has a high specificity for phosphobutidine and a
CC molecular weight of 13000-15000 Da. The histidine phosphatase gene is
CC localised at chromosome 9 (9q33). The histidine phosphatase can be used
CC for diagnosis and treatment of pathological states of cell regulation
CC and cell growth. These include cancers, immune disorders, viral
CC infection, genetic disorders, and heart disease. The histidine
CC phosphatase can also be used for identifying agonists and antagonists
CC which can be used to treat conditions associated with N-phosphorylation
CC imbalance.

SQ Sequence 125 AA;

Query Match	100.0%;	Score 674;	DB 21;	Length 125;
Best Local Similarity	100.0%;	Pred. No. 6.2e-74;		
Matches 125; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 MAVADLALIPDVIDSDGYFKYLLRVHSAPRSGAPAESKEIIRGYNAEYHADIIYDKV 60
Db 1 MAVADLALIPDVIDSDGYFKYLIRVHSAPRSGAPAESKEIIRGYKNAEYHADIYDKV 600

Oy 61 SGDMDKQSGDCECLGGGRISHSODKTHVGYGSMAYGAQHAISTEKIKAKYPDYEVTTW 120
|||
Db 61 SGDMDKQSGDCECLGGGRISHSODKTHVGYGSMAYGAQHAISTEKIKAKYPDYEVTTW 120

Qy	121	ANDGY	125
Db	121	ANDGY	125

RESULT 2
AAM39661
ID AAM39661 standard; Protein; 125 AA

AC AAM39661;

DT 22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 2806.

KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemolactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens

PN W0200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725
PR 25-APR-2000; 2000US-0553317
PR 09-JUL-2000; 2000US-0598042
PR 19-JUL-2000; 2000US-0620312
PR 03-AUG-2000; 2000US-0653450
PR 14-SEP-2000; 2000US-0662191
PR 19-OCT-2000; 2000US-0693036
PR 29-NOV-2000; 2000US-0727344

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47
DR
DR N-PSDB; AAI58817.
XX

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2806; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ Sequence 125 AA;

Query Match	100.0%	Score 674	DB 22	Length 125
Best Local Similarity	100.0%	Pred. No. 6	2e-74	
Matches 125; Conservative	0	Mismatches	0	Indels 0; Gaps 0;

QY 1 MAVADLALIPVDIDSDGVFKYVILIRVHSAPRSGAPAEESKEIVRCYKWAETHADIDYDKV 60

DB I MAVADLALIPDVDDIDSDGVFKYILIRVSHAPRSGBPAAESKEIIVRGINMAEIHADIIDNV 60

QY 61 SGDMQKGCDCCECLGGGRISHSQDCKIHVYGYSMAYGPAQHAISTEKIKAKYPDEVTW 120

Db	61	SGDMQKQGC
QY	121	ANDGY 125

Db 121 ANDGY 125

RESULT 3
AAV48348
ID AAV48348 standard; Protein; 165 AA

AC AY48348;

DT 08-DEC-1999 (first entry)

DE Human prostate cancer-associated protein 45

KM Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
KM cancer; tissue specificity; human.

OS Homo sapiens

PN DE19811194-A1.

PD 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011194.

PR 10-MAR-1998; 98DE-1011194.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinemann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A,
XX
DR WPI; 1999-519629/44.

DR N-PSDB: AA233497.
 XX
 PT New nucleic acid expressed at high level in normal prostatic tissue and
 PT encoded polypeptides, used to treat cancer and screen for therapeutic
 PT agents
 XX
 PS Claim 22; 141; 194pp; German.
 XX
 CC This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)
 CC encoded by (A) are used: (a) for identifying agents for treatment of
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally
 CC where expressed by gene therapy methods. (A) is also used to isolate
 CC full-length genes (for gene therapy) and for recombinant production of
 CC (I), which can be used to raise specific antibodies. (A) are identified
 CC by assembly of ESTs (expressed sequence tags) before these are analyzed
 CC for expression pattern (tissue specificity). This approach eliminates
 CC many of the false results, as regards tissue specificity, associated
 CC with known methods that use single (usually short) ESTs. AA14304-Y48456
 CC represent peptides encoded by the expressed sequence tags described in
 CC the method of the invention.
 XX
 SQ Sequence 165 AA;
 Query Match 100.0%; Score 674; DB 20; Length 165;
 Best Local Similarity 100.0%; Pred. No. 9.1e-74;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAVADLALIPVDIDSDGVEFKYVLIIRVHSAPRSGAPAAESKEIVRGYKMAEYHADIDYKV 60
 Db 41 MAVADLALIPVDIDSDGVEFKYVLIIRVHSAPRSGAPAAESKEIVRGYKMAEYHADIDYKV 100
 QY 61 SGDMQKQCGDCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDEYVTW 120
 Db 101 SGDMQKQCGDCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDEYVTW 160
 QY 121 ANDGY 125
 Db 161 ANDGY 165
 XX
 RESULT 4
 AA41447
 ID AA41447 standard; Protein; 165 AA.
 AC
 XX
 XX
 DT 22-OCT-2001 (first entry)
 DE
 XX
 DE Human polypeptide SEQ ID NO 6378.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW Alzheimertic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 XX
 OS Homo sapiens.
 XX
 PN WO20015312-A1.
 XX
 PD 26-JUL-2001.
 PF
 XX
 XX 26-DEC-2000; 2000MO-US34263.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB: AA160603.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 2; SEQ ID NO 6378; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S. disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 165 AA;
 Query Match 100.0%; Score 674; DB 22; Length 165;
 Best Local Similarity 100.0%; Pred. No. 9.1e-74;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAVADLALIPVDIDSDGVEFKYVLIIRVHSAPRSGAPAAESKEIVRGYKMAEYHADIDYKV 60
 Db 41 MAVADLALIPVDIDSDGVEFKYVLIIRVHSAPRSGAPAAESKEIVRGYKMAEYHADIDYKV 100
 QY 61 SGDMQKQCGDCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDEYVTW 120
 Db 101 SGDMQKQCGDCCECLGGGRISHOSQDKKIHYGYSMAYGPAQHAISTEKIKAKYPDEYVTW 160
 QY 121 ANDGY 125
 Db 161 ANDGY 165
 XX
 RESULT 5
 AAB08791
 ID AAB08791 standard; Peptide; 124 AA.
 XX
 XX
 AC AAB08791;
 XX
 DT 02-JAN-2001 (first entry)
 DE
 XX
 DE Rabbit histidine protein phosphatase polypeptide.
 XX
 KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 KW cell regulation; cell growth; cancer; immune disorder; viral infection;
 KW genetic disorder; heart disease; N-phosphorylation imbalance.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO200052175-A1.
 XX
 PD 08-SEP-2000.
 PF
 PF 02-MAR-2000; 2000MO-EP01774.

XX 04-MAR-1999; 99DE-1009388.
 XX (MERE) MERCK PATENT GMBH.
 XX Klumpp S, Kellner R;
 XX WPI: 2000-572187/53.
 XX
 XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
 XX immune disorders, viral infection, genetic disorders, and heart disease
 XX
 XX Claim 7; Page 18; 39pp; English.
 XX
 XX The present sequence represents a rabbit histidine phosphatase
 XX polypeptide. The human polypeptide has a high specificity for
 XX phosphohistidine and a molecular weight of 13000-15000 Da. The human
 XX histidine phosphatase gene is localised at chromosome 9 (9q33). The
 XX human histidine phosphatase can be used for diagnosis and treatment
 XX of pathological states of cell regulation and cell growth. These
 XX include cancers, immune disorders, viral infection, genetic disorders,
 XX and heart disease. The histidine phosphatase can also be used for
 XX identifying agonists and antagonists which can be used to treat
 XX conditions associated with N-phosphorylation imbalance.

XX Sequence 124 AA;

Query Match 84.9%; Score 572; DB 21; Length 124;
 Best Local Similarity 83.9%; Pred. No. 1.6e-61;
 Matches 104; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 2 AAADALIPVDIDSDGYFKYVILRVHSAPRSGAPAAESKEIVGYKMAEYHADIDYKVS 61
 DB 1 AAAGIAQIPVDIDSDGYFKYVILRVHRAAPPSEAPGESKIDIVGYKMAEYHADIDYKVS 60
 QY 62 GDMQKQDCCECLGGGRISHOSQDKIHVGYSMAYGPAQHAISTEKIKAKYPDYETWMA 121
 DB 61 GELQKRGHDCCLGGGRISHOSQDKIHVGYSMGYGHAOSHSTERIKAKYPDYETWMA 120
 QY 122 NDGY 125
 DB 121 DDGY 124

RESULT 6
 AAB08793
 ID AAB08793 standard; Peptide; 124 AA.

AC AAB08793;
 DT 02-JAN-2001 (first entry)

DE Mouse histidine protein phosphatase polypeptide.

KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 KW cell regulation; cell growth; cancer; immune disorder; viral infection;
 KW genetic disorder; heart disease; N-phosphorylation imbalance.

OS Mus sp.
 PN WO200052175-A1.

PD 08-SEP-2000.

PF 02-MAR-2000; 2000WO-EP01774.

PR 04-MAR-1999; 99DE-1009388.

PA (MERE) MERCK PATENT GMBH.

PI Klumpp S, Kellner R;
 XX

DR WPI: 2000-572187/53.
 XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
 XX immune disorders, viral infection, genetic disorders, and heart disease
 XX
 XX Disclosure; Page 12; 39pp; English.
 XX
 XX The present sequence represents a mouse histidine phosphatase
 XX polypeptide. The human polypeptide has a high specificity for
 XX phosphohistidine and a molecular weight of 13000-15000 Da. The human
 XX histidine phosphatase gene is localised at chromosome 9 (9q33). The
 XX human histidine phosphatase can be used for diagnosis and treatment
 XX of pathological states of cell regulation and cell growth. These
 XX include cancers, immune disorders, viral infection, genetic disorders,
 XX and heart disease. The histidine phosphatase can also be used for
 XX identifying agonists and antagonists which can be used to treat
 XX conditions associated with N-phosphorylation imbalance.

SQ Sequence 124 AA;

Query Match 83.5%; Score 563; DB 21; Length 124;
 Best Local Similarity 84.4%; Pred. No. 2e-60;
 Matches 103; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 4 ADALIPVDIDSDGYFKYVILRVHSAPRSGAPAAESKEIVGYKMAEYHADIDYKVS 63
 DB 3 ADGQIPVDIDSDGYFKYVILRVHIAEPSPAKCKEIVGYKMAEYHADIDYKVS 62
 QY 64 MQRQDCCECLGGGRISHOSQDKIHVGYSMAYGPAQHAISTEKIKAKYPDYETWMA 123
 DB 63 LQNRGYDCCECLGGGRISHOSQDKIHVGYSMGYGHAOSHSTERIKAKYPDYETWMA 122
 QY 124 GY 125
 DB 123 GY 124

RESULT 7
 AAB42776
 ID AAB42776 standard; Protein; 165 AA.

AC AAB42776;
 DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2540 polypeptide sequence SEQ ID NO:5080.

KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KW vulnery; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antineuritic; antithyroid;
 KW antitumorigenic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disease; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCIP; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

OS Homo sapiens.

PN WO200058473-A2.

PD 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US08621.

PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.

05-APR-1999; 99US-0127728.
30-MAR-2000; 2000US-0540763.
(CURA-) CURAGEN CORP.
Shimkets RA, Leach M;
WPI: 2000-602362/57.
N-PSDB; AAC76985.
Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease -
Claim 11; Page 4264-4265; 5507pp; English.
AAC74446 to AAC77606 encode the proteins given in ABA040237 to ABA043397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnerrary; antiproliferative; antiparkinsonian; nootropic; neuroprotective; osteoplastic; anticonvulsant; antiallergic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antineumatic; antihypertensive; antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergic, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.
Sequence 165 AA:
Query Match 75.2%; Score 507; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 2e-53;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAVADALIPDVDDIDSDGVFKYLYLIRVHSAPRGAPAAESKEIVRYKMAETIADYDXY 60
DB 41 MAVADALIPDVDDIDSDGVFKYLYLIRVHSAPRGAPAAESKEIVRYKMAETIADYDXY 100
QY 61 SGMOKOGCDECLGGGRISHOSODKKIHYYGSM 95
DB 101 SGMOKOGCDECLGGGRISHOSODKKIHYYGSM 135
RESULT 8
AAB08792
ID AAB08792 standard; Peptide; 123 AA.
XX AAB08792;
AC AAB08792;
XX 02-JAN-2001 (first entry)
XX Rat histidine protein phosphatase polypeptide.
XX DE
XX Rat histidine protein phosphatase polypeptide.
XX DE
XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
XX Rattus sp.
XX OS
XX PN WO200052175-A1.
XX PD 08-SEP-2000.

02-MAR-2000; 2000WO-EP01774.
04-MAR-1999; 99DE-009388.
(MERE) MERCK PATENT GMBH.
Klump S, Kellner R;
WPI: 2000-572187/53.
Histidine phosphatase, useful for diagnosis and treatment of cancers, immune disorders, viral infection, genetic disorders, and heart disease
Disclosure; Page 12; 39pp; English.
The present sequence represents a rat histidine phosphatase polypeptide. The human polypeptide has a high specificity for phosphohistidine and a molecular weight of 13000-15000 Da. The human histidine phosphatase gene is localised at chromosome 9 (9q33). The human histidine phosphatase can be used for diagnosis and treatment of pathological states of cell regulation and cell growth. These include cancers, immune disorders, viral infection, genetic disorders, and heart disease. The histidine phosphatase can also be used for identifying agonists and antagonists which can be used to treat conditions associated with N-phosphorylation imbalance.
Sequence 123 AA:
Query Match 52.4%; Score 353; DB 21; Length 123;
Best Local Similarity 86.1%; Pred. No. 7.4e-35;
Matches 62; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 54 ADIYKVSMDKOGCDECLGGGRISHOSODKKIHVGYSMAYGPAQHAISTEKIRAKY 113
DB 52 ADIYKVSMDKOGCDECLGGGRISHOSODKKIHVGYSMAYGPAQHAISTEKIRAKY 111
QY 114 PDYEVTVANDGY 125
DB 112 PDYEVTVANDGY 123
RESULT 9
AAU031440
ID AAU031440 standard; Protein; 119 AA.
XX AAU031440;
XX 18-DEC-2001 (first entry)
XX DE
XX Novel human secreted protein #1931.
XX DE
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
XX OS
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX 16-APR-2001; 2001WO-US08656.
XX PF
XX 18-APR-2000; 2000US-0552929.
XX PR 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
XX PA
XX Tang YT, Liu C, Drmanac RT;
XX WPI: 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 440; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU9510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 CC
 XX
 SQ Sequence 119 AA;
 Query Match 40.6%; Score 273.5; DB 22; Length 119;
 Best Local Similarity 64.0%; Pred. No. 3.4e-25;
 Matches 64; Conservative 4; Mismatches 27; Indels 5; Gaps 4;
 QY 1 MAVADLALIPVDIDSGVFYVLI-RVHSAVRSAGAPAAESKEIVRGYKMAEYHADIDYK 59
 DB 8 MAVADLIPVDIDSGVFYVLIIPSPGIPAPGIRAEKEIVRGYKMAEYHADIDYK 67
 QY 60 -VSGDMQKGC-DCECGGGRISHQSG--DKIHVYGSM 95
 DB 68 SVGATCRKQGLRVISLGGGRISHQSGPQGERFTVYGISM 107
 RESULT 10
 AAB08790
 ID AAB08790 standard; Peptide: 44 AA.
 AC AAB08790;
 XX
 DT 02-JAN-2001 (first entry)
 XX
 DE Conserved sequence of mammalian histidine protein phosphatases.
 XX
 KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 KW cell regulation; cell growth; cancer; immune disorder; viral infection;
 KW genetic disorder; heart disease; N-phosphorylation imbalance.
 XX
 OS Mammalia sp.
 XX
 PN WO200052175-A1.
 PD 08-SEP-2000.
 XX
 PE 02-MAR-2000; 2000WO-EP01774.
 XX
 PR 04-MAR-1999; 99DE-1009388.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Klumpp S, Kellner R;
 XX
 DR WPI; 2000-572187/53.
 XX
 PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
 PT immune disorders, viral infection, genetic disorders, and heart disease
 PT
 PS Claim 4; Page 17; 39pp; English.

XX
 CC The present sequence represents a conserved sequence found in mammalian
 CC histidine phosphatase polypeptides. The human polypeptide has a high
 CC specificity for phosphohistidine and a molecular weight of 13000-15000
 CC Da. The human histidine phosphatase gene is localised at chromosome 9
 CC (9q33). The human histidine phosphatase can be used for diagnosis and
 CC treatment of pathological states of cell regulation and cell growth.
 CC These include cancers, immune disorders, viral infection, genetic
 CC disorders, and heart disease. The histidine phosphatase can also be
 CC used for identifying agonists and antagonists which can be used to
 CC treat conditions associated with N-phosphorylation imbalance.
 CC
 XX
 SQ Sequence 44 AA;
 Query Match 36.9%; Score 249; DB 21; Length 44;
 Best Local Similarity 100.0%; Pred. No. 8.2e-23;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 52 YHADIDKYSQDMQKGCCECGGGRISHQSGDKKIHVYGSM 95
 DB 1 YHADIDKYSQDMQKGCCECGGGRISHQSGDKKIHVYGSM 44
 RESULT 11
 ABB58764
 ID ABB58764 standard; Protein; 135 AA.
 AC ABB58764;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 3084.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02867.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 3084; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pt_sequences.
 CC
 XX
 SQ Sequence 135 AA;

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:25 ; Search time 8.77982 Seconds
(Without alignments)
110.590 Million cell updates/sec

Title: US-09-914-831-4

Sequence: 1 DCECLGGGRSHSQDKIHVGYSMXGYAQR 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/6C_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48.5	27.7	388	1	US-08-290-448A-80
2	48.5	27.7	388	1	US-08-290-448A-80
3	48.5	27.7	388	1	US-08-175-069A-80
4	48.5	27.7	388	4	US-08-461-939B-80
5	48.5	27.7	388	4	US-08-464-000-80
6	46	26.3	48	4	US-08-438-745-22
7	46	26.3	48	5	PCR-US94-0569A-22
8	45.5	26.0	1007	2	US-08-551-459-4
9	45	25.7	503	4	US-09-562-737-65
10	44.5	25.4	303	2	US-08-599-171A-32
11	44.5	25.4	303	2	US-08-646-590B-32
12	44.5	25.4	303	3	US-09-069-226-32
13	44.5	25.4	303	4	US-09-412-184-32
14	44	25.1	353	4	US-09-171-156A-65
15	44	25.1	353	4	US-09-171-156A-65
16	44	25.1	353	4	US-08-697-954-4
17	44	25.1	449	4	US-08-697-954-4
18	44	25.1	503	4	US-09-562-737-67
19	44	25.1	548	4	US-09-167-299-3
20	44	25.1	1264	1	US-07-789-915A-6
21	44	25.1	1264	1	US-08-005-002C-6
22	43.5	24.9	376	2	US-08-487-203A-6
23	43.5	24.9	389	2	US-08-485-449-5
24	43.5	24.9	389	2	US-08-485-449-5
25	43.5	24.9	389	2	US-08-485-449-6
26	43.5	24.9	911	1	US-08-596-985-2
27	43	24.6	352	1	US-07-923-260A-2

28	43	24.6	352	1	US-07-923-260A-3	Sequence 3, Appl
29	43	24.6	381	1	US-07-772-087-6	Sequence 6, Appl
30	43	24.6	381	1	US-08-173-508-12	Sequence 12, Appl
31	43	24.6	381	2	US-08-265-310-12	Sequence 12, Appl
32	43	24.6	382	3	US-08-951-742-12	Sequence 2, Appl
33	43	24.6	382	1	US-08-460-343B-2	Sequence 2, Appl
34	43	24.6	382	1	US-08-398-028B-2	Sequence 2, Appl
35	43	24.6	382	2	US-08-504-265B-2	Sequence 2, Appl
36	43	24.6	382	4	US-09-255-502-2	Sequence 2, Appl
37	43	24.6	382	4	US-09-178-155-2	Sequence 2, Appl
38	43	24.6	382	4	US-09-445-472-31	Sequence 31, Appl
39	43	24.6	382	6	5472855-2	Patent No. 5472855
40	43	24.6	592	3	US-08-991-813-2	Sequence 2, Appl
41	43	24.6	1329	1	US-08-100-709-4	Sequence 4, Appl
42	43	24.6	1329	1	US-08-176-865-4	Sequence 4, Appl
43	43	24.6	1329	1	US-08-474-038-4	Sequence 4, Appl
44	43	24.6	1329	2	US-08-779-046-4	Sequence 4, Appl
45	43	24.6	1329	2	US-08-881-340-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-290-448A-80
Sequence 80, Application US/08290448A
Patent No. 5676954
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rainier, Thonum
APPLICANT: Kuo, Mel-chang
TITLE OF INVENTION: Allergenic Proteins From Rayweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-448A-80

Query Match 27.7%; Score 48.5; DB 1; Length 388;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

OY 1 DCECLGGGRIS-----HOSODXKIHVYGY 25
| | | | |
Db 162 DVRVLPGRIRKSNNGPAIRPHQSDGDAIHVTGSS 195

RESULT 2

US-08-290-448A-80
; Sequence 80, Application US/08290448A
; Patent No. 5698204

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Ratner, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-448A-80

Query Match 27.7%; Score 48.5; DB 1; Length 388;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

OY 1 DCECLGGGRIS-----HOSODXKIHVYGY 25
| | | | |
Db 162 DVRVLPGRIRKSNNGPAIRPHQSDGDAIHVTGSS 195

RESULT 3

US-08-175-069A-80
; Sequence 80, Application US/08175069A
; Patent No. 5776761

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Ratner, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-175-069A-80

Query Match 27.7%; Score 48.5; DB 1; Length 388;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

OY 1 DCECLGGGRIS-----HOSODXKIHVYGY 25
| | | | |
Db 162 DVRVLPGRIRKSNNGPAIRPHQSDGDAIHVTGSS 195

RESULT 4

US-08-461-939B-80
; Sequence 80, Application US/08461939B
; Patent No. 6335019

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Ratner, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Methods For Treating Sensitivity To A
TITLE OF INVENTION: Protein Allergen Using Peptides Which Include A T Cell Epit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,939B
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/464,000
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/290,448

FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-939B-80

Query Match 27.7%; Score 48.5; DB 4; Length 388;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

QY 1 DCECLGGGRIS-----HOSQDKIHVYGS 25
DB 162 DVAVLPFGRIKSNNGPAIPRHQSDDAIHVTGSS 195

RESULT 5
US-08-464-000-80
Sequence 80, Application US/08464000
Patent No. 6335020
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Ratnar, Thorunn
APPLICANT: Kuo, Mel-chang
TITLE OF INVENTION: Allergenic Peptides from Ragweed Pollen
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,000
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/290,448
FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-464-000-80

Query Match 27.7%; Score 48.5; DB 4; Length 388;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

QY 1 DCECLGGGRIS-----HOSQDKIHVYGS 25
DB 162 DVAVLPFGRIKSNNGPAIPRHQSDDAIHVTGSS 195

RESULT 6
US-08-438-745-22
Sequence 22, Application US/08438745
Patent No. 6248715
GENERAL INFORMATION:
APPLICANT: Rosenberg, Steven
APPLICANT: Stratton-Thomas, Jennifer
TITLE OF INVENTION: Expression of Urokinase Plasminogen
Activator Inhibitors
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,745
FILING DATE: 10-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/070,153
FILING DATE: 01-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-438-745-22

Query Match 26.3%; Score 46; DB 4; Length 48;
Best Local Similarity 53.8%; Pred. No. 4.1;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHQ 13
DB 10 NCDCLNGTCVHQ 22

RESULT 7

PCT-US94-05669A-22
Sequence 22, Application PC/TUS94/05669A
GENERAL INFORMATION:
APPLICANT: Chiron Corporation
TITLE OF INVENTION: Expression of Urokinase Plasminogen
NUMBER OF SEQUENCES: 22
TITLE OF INVENTION: Activator Inhibitors
CORRESPONDENCE ADDRESS:
ADDRESS: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05669A
FILING DATE: 19-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939.100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05669A-22
Query Match
Best Local Similarity 53.8%; Score 46; DB 5; Length 48;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 1 DCECLGGGRISHQ 13
Db 10 NCDCLNGTCVHQ 22
RESULT 8.
US-08-551-459-4
Sequence 4, Application US/08551459
Patent No. 5821350
GENERAL INFORMATION:
APPLICANT: Huang, Yue
APPLICANT: Karatzas, Costas N.
APPLICANT: Lazaris-Karlzas, Anthoula
APPLICANT: Delaquis, Annick
TITLE OF INVENTION: ASPERGILLUS NIGER BETA GALACTOSIDASE
TITLE OF INVENTION: GENE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/551,459

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06632/005001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1007 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-551-459-4
Query Match
Best Local Similarity 32.3%; Score 45.5; DB 2; Length 1007;
Matches 10; Conservative 6; Mismatches 10; Indels 5; Gaps 1;
OY 7 GGRISHOSQDKKHVGYGSMK-----YGAQ 32
Db 460 GGTTLNGRDSKIHVDYNSGNIITSTAE 490
RESULT 9
US-09-562-737-65
Sequence 65, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 65
LENGTH: 503
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-65
Query Match
Best Local Similarity 61.58%; Score 45; DB 4; Length 503;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 10 ISHOSQDKKHVY 22
Db 114 VTHDSQDKKIFY 126
RESULT 10
US-08-599-171A-32
Sequence 32, Application US/08599171A
Patent No. 5814473
GENERAL INFORMATION:
APPLICANT: WARREN, Patrick V.
TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESS: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA

ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,171A
FILING DATE: Concurrently
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HERRON, CHARLES J.
REGISTRATION NUMBER: 28,019
REFERENCE/DOCKET NUMBER: 331400-38
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-599-171A-32

Query Match 25.4%; Score 44.5; DB 2; Length 303;
Best Local Similarity 41.7%; Pred. No. 53;
Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 8 GRISHOSQDKIHVGYSMYXGA 31
||| : : |||| : : || :
Db 12 GRIL-KWEDAKIHVLTALHYGTS 34

RESULT 11
US-08-646-590B-32
Sequence 32, Application US/08646590B
Patent No. 5962283
GENERAL INFORMATION:
APPLICANT: Warren, Patrick V.
APPLICANT: Swanson, Ronald V.
TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,590B
FILING DATE: 08-May-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,171
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01094
FILING DATE: 21-January-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 09010/017001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-646-590B-32

Query Match 25.4%; Score 44.5; DB 2; Length 303;
Best Local Similarity 41.7%; Pred. No. 53;
Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 8 GRISHOSQDKIHVGYSMYXGA 31
||| : : |||| : : || :
Db 12 GRIL-KWEDAKIHVLTALHYGTS 34

RESULT 12
US-09-069-226-32
Sequence 32, Application US/09069226
Patent No. 6013509
GENERAL INFORMATION:
APPLICANT: WARREN, Patrick V.
TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/069,226
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,171
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HERRON, CHARLES J.
REGISTRATION NUMBER: 28,019
REFERENCE/DOCKET NUMBER: 331400-38
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-09-069-226-32

Query Match 25.4%; Score 44.5; DB 3; Length 303;
Best Local Similarity 41.7%; Pred. No. 53;
Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 8 GRISHOSQDKIHVGYSMYXGA 31
||| : : |||| : : || :
Db 12 GRIL-KWEDAKIHVLTALHYGTS 34

RESULT 13
US-09-412-184-32
Sequence 32, Application US/09412184
Patent No. 6268188
GENERAL INFORMATION:
APPLICANT: Warren, Patrick V.
APPLICANT: Swanson, Ronald V.
TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/412,184
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,590
FILING DATE: 08-May-1996
APPLICATION NUMBER: 08/599,171
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01094
FILING DATE: 21-January-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/017001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-412-184-32
Query Match 25.4%; Score 44.5; DB 4; Length 303;
Best Local Similarity 41.7%; Pred. No. 53;
Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;
QY 8 GRISHOSDKKHVGYGSMYGA 31
DB 12 GRIL-KWEDAKIHVTHALHYGTS 34
RESULT 14
US-09-171-156A-65
Sequence 65, Application US/09171156A
Patent No. 6368846
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
Sim, Gek-Kee
Weber, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223

STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
US-09-171-156A-65
Query Match 25.1%; Score 44; DB 4; Length 353;
Best Local Similarity 46.7%; Pred. No. 74;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 2 CECGGGRISHOSD 16
DB 17 CDYGGKTHKSD 31
RESULT 15
US-09-171-156A-68
Sequence 68, Application US/09171156A
Patent No. 6368846
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
Sim, Gek-Kee
Weber, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223

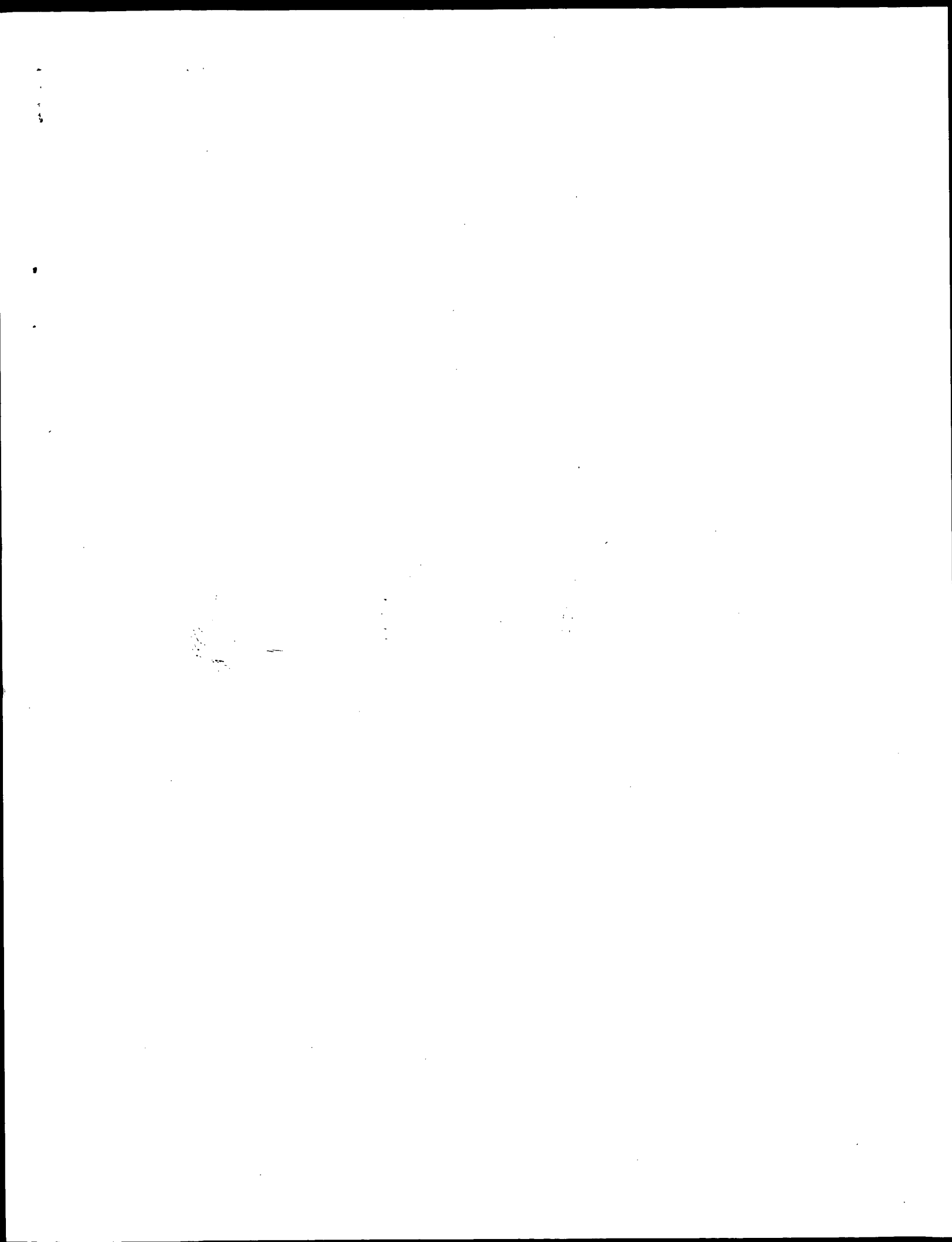
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-171-156A-68

Query Match 25.1%; Score 44; DB 4; Length 353;
Best Local Similarity 46.7%; Pred. No. 74;

Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 CECLEGGGRISHQSD 16
I: I: I: I: I: I:
Db 17 CDYPGGPKITHKSD 31

Search completed: May 7, 2003, 19:08:37
Job time : 9.77982 secs



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:00:25 ; Search time 22.8578 Seconds
(Without alignments)
297,472 Million cell updates/sec

Title: US-09-914-831-4

Perfect score: 175
Sequence: 1 DCECLGGGRISHQSDPKIHVYGYGSMYXGAOH 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP_invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP-organelle:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP_unclassified:*
15: SP-rvivirus:*
16: SP_bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169	96.6	124	11 Q9DAK9	Q9DAK9 mus musculus
2	169	96.6	125	4 Q9NRX4	Q9NRX4 homo sapien
3	164	93.7	125	4 Q9H0Y3	Q9H0Y3 homo sapien
4	91	52.0	115	5 P90861	P90861 caenorhabdi
5	91	52.0	135	5 Q9BH48	Q9BH48 drosophila
6	89	50.9	102	5 Q9BH68	Q9BH68 drosophila
7	89	50.9	102	5 Q9BH67	Q9BH67 drosophila
8	89	50.9	104	5 Q9BM25	Q9BM25 drosophila
9	89	50.9	104	5 Q9BM21	Q9BM21 drosophila
10	89	50.9	109	5 Q9BM27	Q9BM27 drosophila
11	89	50.9	119	5 Q9BM475	Q9BM475 drosophila
12	89	50.9	135	5 Q9BM98	Q9BM98 drosophila
13	89	50.9	135	5 Q9BM97	Q9BM97 drosophila
14	89	50.9	135	5 Q9BM96	Q9BM96 drosophila
15	89	50.9	135	5 Q9BM99	Q9BM99 drosophila
16	89	50.9	135	5 Q95VY7	Q95VY7 drosophila

17	78	44.6	140	5 Q9BM91	Q9BM91 drosophila
18	76	43.4	140	5 Q9BM94	Q9BM94 drosophila
19	76	43.4	140	5 Q9BM93	Q9BM93 drosophila
20	76	43.4	140	5 Q9BM95	Q9BM95 drosophila
21	75	42.9	140	5 Q9BM92	Q9BM92 drosophila
22	75	42.9	140	5 Q9BM90	Q9BM90 drosophila
23	75	42.9	140	5 Q9BM89	Q9BM89 drosophila
24	68	38.9	140	5 Q9BM85	Q9BM85 drosophila
25	68	38.9	140	5 Q9BM84	Q9BM84 drosophila
26	68	38.9	140	5 Q9BM82	Q9BM82 drosophila
27	66	37.7	140	5 Q9BM87	Q9BM87 drosophila
28	66	37.7	140	5 Q9BM86	Q9BM86 drosophila
29	66	37.7	140	5 Q9BM83	Q9BM83 drosophila
30	66	37.7	140	5 Q9BM88	Q9BM88 drosophila
31	66	37.7	148	5 Q9Y170	Q9Y170 drosophila
32	61	34.9	160	5 Q9V1G0	Q9V1G0 drosophila
33	57	32.6	161	5 Q8T2P4	Q8T2P4 dictyosteli
34	55	31.4	494	3 Q9C1E2	Q9C1E2 botrytis ci
35	55	31.4	496	3 Q9HDR0	Q9HDR0 botrytis ci
36	55	31.4	498	3 Q9HDR6	Q9HDR6 botrytis ci
37	55	31.4	498	3 Q9HDR1	Q9HDR1 botrytis ci
38	55	31.4	498	3 Q9C1E3	Q9C1E3 botrytis ci
39	55	31.4	500	3 Q9HDR5	Q9HDR5 botrytis ci
40	55	31.4	500	3 Q9C1E5	Q9C1E5 botrytis ci
41	55	31.4	501	3 Q9HDR4	Q9HDR4 botrytis ci
42	55	31.4	501	3 Q9HDR3	Q9HDR3 botrytis ci
43	55	31.4	501	3 Q9HDP9	Q9HDP9 botrytis ci
44	55	31.4	501	3 Q9C1E4	Q9C1E4 botrytis ci
45	55	31.4	502	3 Q9C1E1	Q9C1E1 botrytis ci

ALIGNMENTS

RESULT 1
Q9DAK9 PRELIMINARY; PRT; 124 AA.
AC Q9DAK9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 1700008C22R1k protein (RIKEN cDNA I700008C22 gene).
GN 1700008C22R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Furuta S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Felschmann W., Gaasterland T., Gissi C., King B., Kocha H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barz G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE-MAMMARY GLAND:
 RA Strausberg R.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK005756; BAB24222.1; -
 DR EMBL: BC028657; AAH28657.1; -
 DR MGD: MGI:1922704; 1700008C22RLK.
 SQ SEQUENCE 124 AA; 13996 MW; 589FA3CE95A2F66 CRC64;

Query Match 96.6%; Score 169; DB 11; Length 124;
 Best Local Similarity 90.9%; Pred. No. 4.7e-17;
 Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 DCEICGGGRISHOSODKKIHVGYGSMYXGAOH 33
 DB 69 DCEICGGGRISHOSODKKIHVGYGSMYXGAOH 101

RESULT 2

ID O9NRX4 PRELIMINARY; PRT; 125 AA.

AC O9NRX4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Sex-regulated protein janus-A (CGI-202) (Similar to HSPC141 protein).

OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE-ADRENAL GLAND;
 RA Li Y., Shi J., Huang C., Jiang C., Ren S., Zhou J., Yu Y., Xu S.,
 RA Wang Y., Fu G., Chen Z., Han Z.;

RT "A novel gene expressed in human adrenal gland."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-ADRENAL GLAND;
 RA Peng Y., Gu Y., Li Y., Fu S., Gu J., Gu W., Jiang C., Yu Y., Han Z.,
 RA Wang Y., Chen Z., Fu G.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.

RA Lin W.-C.;

RT "Identification of novel human genes by comparative proteomics."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RA Strausberg R.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF164793; AAF80759.1; -
 DR EMBL: AF285119; AAG01156.1; -
 DR EMBL: BC024648; AAH24648.1; -

SQ SEQUENCE 125 AA; 13832 MW; 24FOCA2BADB78478 CRC64;

Query Match 96.6%; Score 169; DB 4; Length 125;
 Best Local Similarity 90.9%; Pred. No. 4.8e-17;
 Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 DCEICGGGRISHOSODKKIHVGYGSMYXGAOH 33
 DB 70 DCEICGGGRISHOSODKKIHVGYGSMYXGAOH 102

RESULT 3

ID O9HOY3 PRELIMINARY; PRT; 125 AA.

AC O9HOY3;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 13.8 kDa protein.
 GN DKEP564M173.

OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RA Strausberg R.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF164793; AAF80759.1; -
 DR EMBL: AF285119; AAG01156.1; -
 DR EMBL: BC024648; AAH24648.1; -

SQ SEQUENCE 125 AA; 13822 MW; 3C475F5B43ADE5A CRC64;

Query Match 93.7%; Score 164; DB 4; Length 125;
 Best Local Similarity 87.9%; Pred. No. 2.5e-16;
 Matches 29; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 DCEICGGGRISHOSODKKIHVGYGSMYXGAOH 33
 DB 70 DCEICGGGRISHOSODKKIHVGYGSMYXGAOH 102

RESULT 4

ID P90861 PRELIMINARY; PRT; 115 AA.

AC P90861;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F36A2.8 protein.
 GN F36A2.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.
 RA Lennard N.;

RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RA Strausberg R.;

RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF164793; AAF80759.1; -
 DR EMBL: AF285119; AAG01156.1; -
 DR EMBL: BC024648; AAH24648.1; -

SQ SEQUENCE 115 AA; 13183 MW; 29BC8720EC09203 CRC64;

Query Match 52.0%; Score 91; DB 5; Length 115;
 Best Local Similarity 54.8%; Pred. No. 1e-05;
 Matches 17; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

OY 3 EICIGGGGRISHOSODKKIHVGYGSMYXGAOH 33
 DB 62 KCYGGGRIRHDELGRDILVGYGYGRADH 92

RESULT 5

ID O9BH48 PRELIMINARY; PRT; 135 AA.

AC O9BH48;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)

RESULT 7
Q9BH67
ID Q9BH67 PRELIMINARY; PRT; 102 AA

```

QY      3 ECLGGGRISHQSDKIIHYGYSMYXGAQH 33
          || ||||| : ||||| : |||
Db      51 ECPGGRIEHNPDKKYLKVGYSGFGKADH 81

```

RESULT 9

Q9BM21 PRELIMINARY; PRT; 104 AA.

ID Q9BM21
AC Q9BM21
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Janus A (Fragment).
GN JANA.

OS Drosophila sechellia (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7238;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=228;
RX MEDLINE=20556156; PubMed=11102384;
RA Kliman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
RA Berry A.J., McCarter J., Wakeley J., Hey J.;
RT "The Population Genetics of the Origin and Divergence of the
RT Drosophila simulans Complex Species.";
RL Genetics 156:1913-1931(2000).
DR EMBL: AF284459; AAC49478.1; -.
DR FlyBase; FBgn0043637; DsecJana.
FT NON_TER
SQ SEQUENCE 104 AA; 11735 MW; 73BF32548B261427 CRC64;

Query Match 50.9%; Score 89; DB 5; Length 104;
Best Local Similarity 51.6%; Pred. No. 1.8e-05;
Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHOSQDKIHVGYSMXYGAQH 33
DB 51 ECPGGRIENHPDKKYLKVGYSQGFGRADH 81

RESULT 10
Q9BM27 PRELIMINARY; PRT; 109 AA.

ID Q9BM27
AC Q9BM27
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Janus A (Fragment).
GN JANA.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=KENYA.2;
RX MEDLINE=20556156; PubMed=11102384;
RA Kliman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
RA Berry A.J., McCarter J., Wakeley J., Hey J.;
RT "The Population Genetics of the Origin and Divergence of the
RT Drosophila simulans Complex Species.";
RL Genetics 156:1913-1931(2000).
DR EMBL: AF284453; AAC49466.1; -.
DR FlyBase; FBgn0016348; DsimJana.
FT NON_TER
SQ SEQUENCE 109 AA; 12384 MW; 1118CE3722E56A6F CRC64;

Query Match 50.9%; Score 89; DB 5; Length 109;
Best Local Similarity 51.6%; Pred. No. 1.9e-05;
Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHOSQDKIHVGYSMXYGAQH 33
DB 56 ECPGGRIENHPDKKYLKVGYSQGFGRADH 86

RESULT 11
Q8T475 PRELIMINARY; PRT; 119 AA.

ID Q8T475
AC Q8T475
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE AT12574P.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Lao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paclebo J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celiker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY089320; AAL90058.1; -.
DR FlyBase; FBgn0043691; DmanJana.
FT NON_TER
SQ SEQUENCE 119 AA; 13300 MW; 3EFF85BCE91F8068 CRC64;

Query Match 50.9%; Score 89; DB 5; Length 119;
Best Local Similarity 51.6%; Pred. No. 2.1e-05;
Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHOSQDKIHVGYSMXYGAQH 33
DB 66 ECPGGRIENHPDKKYLKVGYSQGFGRADH 96

RESULT 12
Q9BM98 PRELIMINARY; PRT; 135 AA.

ID Q9BM98
AC Q9BM98
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Janus A.
GN JANA.
OS Drosophila mauritiana (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7226;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=KENYA.2;
RX MEDLINE=21219192; PubMed=11319264;
RA Parsch J., Melkielejohn C.D., Hauschteck-Jungen E., Hunziker P.,
RA Hartl D.L.;
RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
RT melanogaster Species Subgroup.";
RL Mol. Biol. Evol. 18:801-811(2001).
DR EMBL: AY013340; AAC50361.1; -.
DR FlyBase; FBgn0043691; DmanJana.
FT NON_TER
SQ SEQUENCE 135 AA; 15158 MW; 2F91E387BD70C1CC CRC64;

Query Match 50.9%; Score 89; DB 5; Length 135;
Best Local Similarity 51.6%; Pred. No. 2.4e-05;
Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHOSQDKIHVGYSMXYGAQH 33
DB 82 ECPGGRIENHPDKKYLKVGYSQGFGRADH 112

RESULT 13
Q9BM97 PRELIMINARY; PRT; 135 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janus.
 GN JANA.
 OS Drosophila erecta (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Meiklejohn C.D., Hauschteck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus genes in the Drosophila
 RT melanogaster Species Subgroup."
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL; AY013343; AAG50364.1; -
 DR FlyBase; FBgn0043709; Dore\Jana.
 SQ SEQUENCE 135 AA; 15180 MW; DB294031FEDEFD07C CRC64;

Query Match 50.9%; Score 89; DB 5; Length 135;
 Best Local Similarity 51.6%; Pred. No. 2.4e-05;
 Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY 3 ECLGGGRISHQSODXKIHVGYSMYGYXAOH 33
 DB 82 ECPGGRIEHPDKKYLKVGYSOGFGKADH 112

RESULT 14

O9BM96 PRELIMINARY; PRT; 135 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janus.
 GN JANA.

OS Drosophila orena (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Meiklejohn C.D., Hauschteck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus genes in the Drosophila
 RT melanogaster Species Subgroup."
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL; AY013344; AAG50365.1; -
 DR FlyBase; FBgn0043667; Dore\Jana.
 SQ SEQUENCE 135 AA; 15146 MW; DB21E237BCD1307C CRC64;

Query Match 50.9%; Score 89; DB 5; Length 135;
 Best Local Similarity 51.6%; Pred. No. 2.4e-05;
 Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY 3 ECLGGGRISHQSODXKIHVGYSMYGYXAOH 33
 DB 82 ECPGGRIEHPDKKYLKVGYSOGFGKADH 112

RESULT 15

O9BM99 PRELIMINARY; PRT; 135 AA.

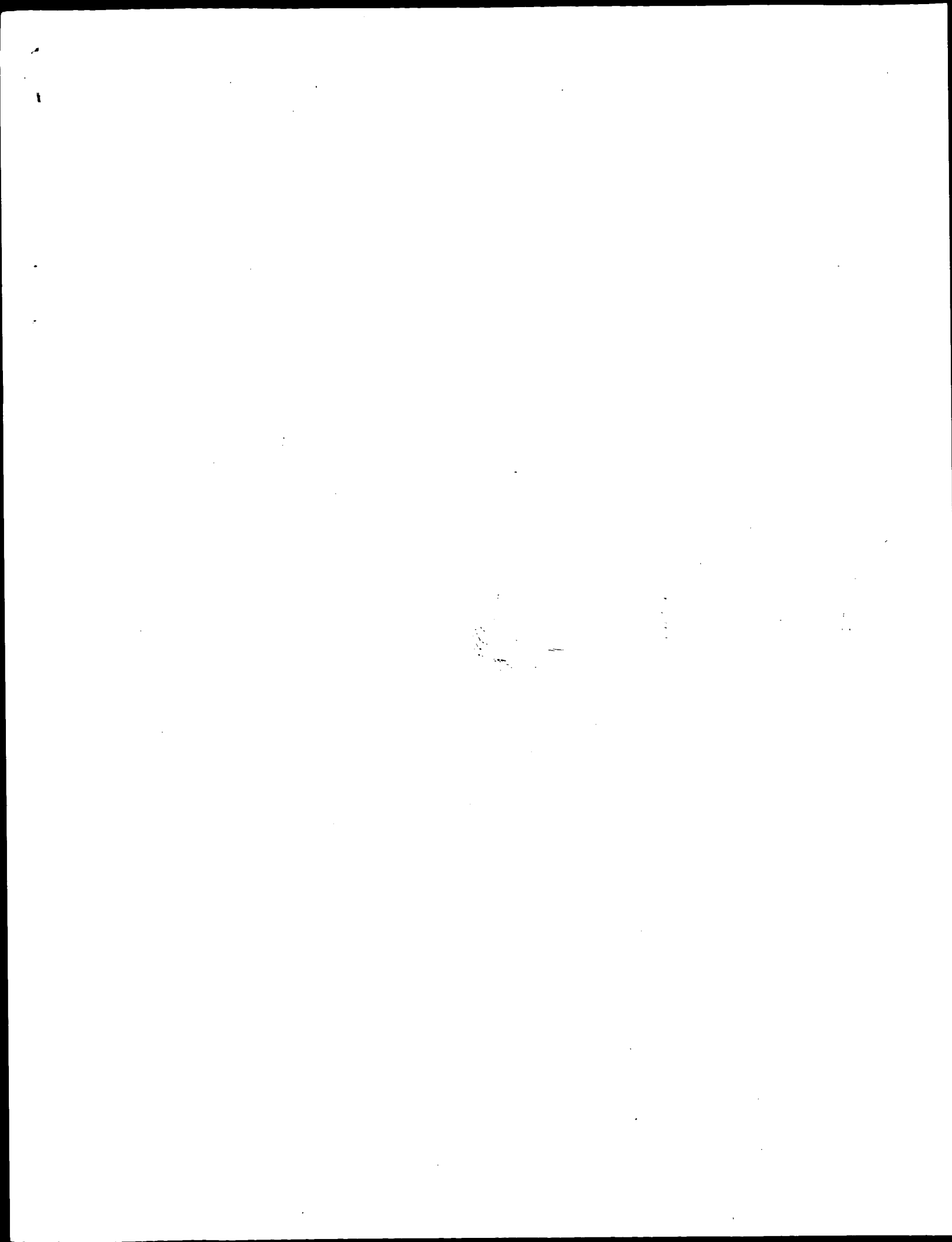
DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Janus.
 GN JANA.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Meiklejohn C.D., Hauschteck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus genes in the Drosophila
 RT melanogaster Species Subgroup."
 RL Mol. Biol. Evol. 18:801-811(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRALIN=536, S3, S4, S5, S6, S7, S8, S17, S19, S25, S31, AND S34;
 RX MEDLINE=21518596; PubMed=11606541;
 RA Parsch J., Meiklejohn C.D., Hartl D.L.;
 RT "Patterns of DNA Sequence Variation Suggest the Recent Action of
 RT Positive Selection in the janus-ocnus Region of Drosophila simulans."
 RL Genetics 159:647-657(2001).
 DR EMBL; AY013339; AAG50360.1; -
 DR EMBL; AF393342; AAK72020.1; -
 DR EMBL; AF393331; AAK72009.1; -
 DR EMBL; AF393332; AAK72010.1; -
 DR EMBL; AF393333; AAK72011.1; -
 DR EMBL; AF393334; AAK72012.1; -
 DR EMBL; AF393335; AAK72013.1; -
 DR EMBL; AF393336; AAK72014.1; -
 DR EMBL; AF393337; AAK72015.1; -
 DR EMBL; AF393338; AAK72016.1; -
 DR EMBL; AF393339; AAK72017.1; -
 DR EMBL; AF393340; AAK72018.1; -
 DR EMBL; AF393341; AAK72019.1; -
 DR FlyBase; FBgn0016348; Dsim\Jana.
 SQ SEQUENCE 135 AA; 15206 MW; 2E2940304F7E21CD CRC64;

Query Match 50.9%; Score 89; DB 5; Length 135;
 Best Local Similarity 51.6%; Pred. No. 2.4e-05;
 Matches 16; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY 3 ECLGGGRISHQSODXKIHVGYSMYGYXAOH 33
 DB 82 ECPGGRIEHPDKKYLKVGYSOGFGKADH 112

Search completed: May 7, 2003, 19:06:38
 Job time : 22.8578 secs



XX Klump S, Kellner R;
XX WPI; 2000-572187/53.
DR Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
PT
XX
PS Claim 3; Page 17; 39pp; English.
CC The present sequence represents a conserved sequence found in mammalian
CC histidine phosphatase polypeptides. The human polypeptide has a high
CC specificity for phosphohistidine and a molecular weight of 13000-15000
CC Da. The human histidine phosphatase gene is localised at chromosome 9
CC (9q33). The human histidine phosphatase can be used for diagnosis and
CC treatment of pathological states of cell regulation and cell growth.
CC These include cancers, immune disorders, viral infection, genetic
CC disorders, and heart disease. The histidine phosphatase can also be
CC used for identifying agonists and antagonists which can be used to
CC treat conditions associated with N-phosphorylation imbalance.
SQ Sequence 33 AA;
Query Match 96.6%; Score 169; DB 21; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DCEICGGGRISHOSQDKRHVGYGSMXYGXAQH 33
DB 1 DCEICGGGRISHOSQDKRHVGYGSMXYGXAQH 33
RESULT 2
ID AAB08792 standard; Peptide: 123 AA.
XX AAB08792;
AC 02-JAN-2001 (first entry)
DT
XX
DE Rat histidine protein phosphatase polypeptide.
XX
XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KM cell regulation; cell growth; cancer; immune disorder; viral infection;
KM genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Rattus sp.
XX
XX WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
XX 04-MAR-1999; 99DE-1009388.
XX
XX (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;
XX
XX WPI; 2000-572187/53.
DR Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
PT
XX
PS Disclosure; Page 12; 39pp; English.
XX The present sequence represents a rat histidine phosphatase
CC polypeptide. The human polypeptide has a high specificity for
CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
CC histidine phosphatase gene is localised at chromosome 9 (9q33). The

CC human histidine phosphatase can be used for diagnosis and treatment
CC of pathological states of cell regulation and cell growth. These
CC include cancers, immune disorders, viral infection, genetic disorders,
CC and heart disease. The histidine phosphatase can also be used for
CC identifying agonists and antagonists which can be used to treat
CC conditions associated with N-phosphorylation imbalance.
SQ Sequence 123 AA;
Query Match 96.6%; Score 169; DB 21; Length 123;
Best Local Similarity 90.9%; Pred. No. 4.5e-16;
Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DCEICGGGRISHOSQDKRHVGYGSMXYGXAQH 33
DB 68 DCEICGGGRISHOSQDKRHVGYGSMXYGXAQH 100
RESULT 3
ID AAB08791 standard; Peptide: 124 AA.
XX AAB08791;
AC 02-JAN-2001 (first entry)
DT
XX
DE Rabbit histidine protein phosphatase polypeptide.
XX
XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KM cell regulation; cell growth; cancer; immune disorder; viral infection;
KM genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Oryctolagus cuniculus.
XX
XX WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
XX 04-MAR-1999; 99DE-1009388.
XX
XX (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;
XX
XX WPI; 2000-572187/53.
DR Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
PT
XX
PS Claim 7; Page 18; 39pp; English.
CC The present sequence represents a rabbit histidine phosphatase
CC polypeptide. The human polypeptide has a high specificity for
CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
CC histidine phosphatase gene is localised at chromosome 9 (9q33). The
CC human histidine phosphatase can be used for diagnosis and treatment
CC of pathological states of cell regulation and cell growth. These
CC include cancers, immune disorders, viral infection, genetic disorders,
CC and heart disease. The histidine phosphatase can also be used for
CC identifying agonists and antagonists which can be used to treat
CC conditions associated with N-phosphorylation imbalance.
SQ Sequence 124 AA;
Query Match 96.6%; Score 169; DB 21; Length 124;
Best Local Similarity 90.9%; Pred. No. 4.5e-16;
Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DCEICGGGRISHOSQDKRHVGYGSMXYGXAQH 33

DB 69 DCECLGGGRISHOSQDKIHVGYGSMGYGPAQH 101

RESULT 4
AAB08793
ID AAB08793 standard; Peptide; 124 AA.
XX
AC AAB08793;
XX
DT 02-JAN-2001 (first entry)
XX
DE Mouse histidine protein phosphatase polypeptide.
XX
KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Mus sp.
XX
PN WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
PR 04-MAR-1999; 99DE-1009388.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;
XX
DR WPI: 2000-572187/53.
XX
PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
XX
PS Disclosure: Page 12; 39pp; English.
XX
CC The present sequence represents a mouse histidine phosphatase
CC polypeptide. The human polypeptide has a high specificity for
CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
CC histidine phosphatase gene is localised at chromosome 9 (9q33). The
CC human histidine phosphatase can be used for diagnosis and treatment
CC of pathological states of cell regulation and cell growth. These
CC include cancers, immune disorders, viral infection, genetic disorders,
CC and heart disease. The histidine phosphatase can also be used for
CC identifying agonists and antagonists which can be used to treat
CC conditions associated with N-phosphorylation imbalance.
XX
SQ Sequence 124 AA:
XX
Query Match 96.6%; Score 169; DB 21; Length 124;
Best Local Similarity 90.9%; Pred. No. 4.5e-16;
Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 DCECLGGGRISHOSQDKIHVGYGSMGYGPAQH 33
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
DB 69 DCECLGGGRISHOSQDKIHVGYGSMGYGPAQH 101
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

RESULT 5
AAB08787
ID AAB08787 standard; Protein; 125 AA.
XX
AC AAB08787;
XX
DT 02-JAN-2001 (first entry)
XX
DE A human histidine protein phosphatase polypeptide.
XX
KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;

KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Homo sapiens.
XX
PN WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
PR 04-MAR-1999; 99DE-1009388.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;
XX
DR WPI: 2000-572187/53.
DR N-PSDB: AAA64502.
XX
PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
XX
PS Claim 5; Page 17; 39pp; English.
XX
CC The present sequence represents human histidine phosphatase polypeptide.
CC The polypeptide has a high specificity for phosphohistidine and a
CC molecular weight of 13000-15000 Da. The histidine phosphatase gene is
CC localised at chromosome 9 (9q33). The histidine phosphatase can be used
CC for diagnosis and treatment of pathological states of cell regulation
CC and cell growth. These include cancers, immune disorders, viral
CC infection, genetic disorders, and heart disease. The histidine
CC phosphatase can also be used for identifying agonists and antagonists
CC which can be used to treat conditions associated with N-phosphorylation
CC imbalance.
XX
SQ Sequence 125 AA:
XX
Query Match 96.6%; Score 169; DB 21; Length 125;
Best Local Similarity 90.9%; Pred. No. 4.6e-16;
Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 DCECLGGGRISHOSQDKIHVGYGSMGYGPAQH 33
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
DB 70 DCECLGGGRISHOSQDKIHVGYGSMGYGPAQH 102
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

RESULT 6
AAM39661
ID AAM39661 standard; Protein; 125 AA.
XX
AC AAM39661;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2806.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB; AAI5817.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 4; SEQ ID NO 2806; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 125 AA;
 XX
 XX
 Query Match 96.6%; Score 169; DB 22; Length 125;
 Best Local Similarity 90.9%; Pred. No. 4.6e-16;
 Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DCECLGGGRISHOSQDKIHVGYGSMXYGXAQH 33
 DB 70 DCECLGGGRISHOSQDKIHVGYGSMAYGPAQH 102
 XX
 RESULT 7
 AAY48348
 ID AAY48348 standard; Protein; 165 AA.
 XX
 AC AAY48348;
 XX
 DT 08-DEC-1999 (first entry)
 XX
 DE Human prostate cancer-associated protein 45.
 XX
 KW Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
 KW cancer; tissue specificity; human.
 XX
 OS Homo sapiens.
 OS
 PN DE19811194-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 10-MAR-1998; 98DE-1011194.
 XX
 PR 10-MAR-1998; 98DE-1011194.
 XX
 XX

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX
 DR WPI: 1999-519629/44.
 DR N-PSDB; AAZ33497.
 XX
 PT New nucleic acid expressed at high level in normal prostatic tissue and
 PT encoded polypeptides, used to treat cancer and screen for therapeutic
 PT agents
 XX
 PS Claim 22; 141; 194pp; German.
 XX
 CC This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)
 CC encoded by (A) are used: (a) for identifying agents for treatment of
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally
 CC where expressed by gene therapy methods. (A) is also used to isolate
 CC full-length genes (for gene therapy) and for recombinant production of
 CC (I), which can be used to raise specific antibodies. (A) are identified
 CC by assembly of ESTs (expressed sequence tags) before these are analyzed
 CC for expression pattern (tissue specificity). This approach eliminates
 CC many of the false results, as regards tissue specificity, associated
 CC with known methods that use single (usually short) ESTs. AA148304-Y48456
 CC represent peptides encoded by the expressed sequence tags described in
 CC the method of the invention.
 CC
 SQ Sequence 165 AA;
 XX
 XX
 Query Match 96.6%; Score 169; DB 20; Length 165;
 Best Local Similarity 90.9%; Pred. No. 6.1e-16;
 Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DCECLGGGRISHOSQDKIHVGYGSMXYGXAQH 33
 DB 110 DCECLGGGRISHOSQDKIHVGYGSMAYGPAQH 142
 XX
 RESULT 8
 AAM41447
 ID AAM41447 standard; Protein; 165 AA.
 XX
 AC AAM41447;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide seq ID NO 6378.
 XX
 KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 OS
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX
 PA (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB: AA160603.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Example 2: SEQ ID NO 6378; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA442213) with nootropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 165 AA;
 XX
 Query Match 96.6%; Score 169; DB 22; Length 165;
 Best Local Similarity 90.9%; Pred. No. 6, 1e-16;
 Matches 30; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 DCECLGGGRISHQSDKKIHVGYSMXGXAQH 33
 DB 110 DCECLGGGRISHQSDKKIHVGYSMXGXAQH 142
 XX
 RESULT 9
 AAB08790
 ID AAB08790 standard; Peptide; 44 AA.
 XX
 AC AAB08790;
 XX
 DT 02-JAN-2001 (first entry)
 XX
 DE Conserved sequence of mammalian histidine protein phosphatases.
 XX
 KW Human: histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 KW cell regulation; cell growth; cancer; immune disorder; viral infection;
 KW genetic disorder; heart disease; N-phosphorylation imbalance.
 XX
 OS Mammalia sp.
 XX
 PN WO200052175-A1.
 XX
 PD 08-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-EP01774.
 XX
 PR 04-MAR-1999; 99DE-1009388.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Klumpp S, Kellner R;
 XX
 DR WPI: 2000-572187/53.
 XX
 PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
 PT immune disorders, viral infection, genetic disorders, and heart disease

PT
 XX
 PS
 XX
 CC Claim 4: Page 17; 39pp; English.
 CC
 CC The present sequence represents a conserved sequence found in mammalian
 CC histidine phosphatase polypeptides. The human polypeptide has a high
 CC specificity for phosphohistidine and a molecular weight of 13000-15000
 CC Da. The human histidine phosphatase gene is localised at chromosome 9
 CC (9q33). The human histidine phosphatase can be used for diagnosis and
 CC treatment of pathological states of cell regulation and cell growth.
 CC These include cancers, immune disorders, viral infection, genetic
 CC disorders, and heart disease. The histidine phosphatase can also be
 CC used for identifying agonists and antagonists which can be used to
 CC treat conditions associated with N-phosphorylation imbalance.
 CC
 SQ Sequence 44 AA;
 XX
 Query Match 80.6%; Score 141; DB 21; Length 44;
 Best Local Similarity 96.2%; Pred. No. 1, 3e-12;
 Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 DCECLGGGRISHQSDKKIHVGYSM 26
 DB 19 DCECLGGGRISHQSDKKIHVGYSM 44
 XX
 RESULT 10
 AAB42776
 ID AAB42776 standard; Protein; 165 AA.
 XX
 AC AAB42776;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2540 polypeptide sequence SEQ ID NO:5080.
 XX
 KW Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vunerary; antiproliferative; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatologic; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-0508621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 XX
 PR 02-APR-1999; 99US-0127636.
 XX
 PR 05-APR-1999; 99US-0127728.
 XX
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA) CURAGEN CORP.
 XX
 PI Shinkens RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 XX
 DR N-PSDB: AAC76985.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,

PT neurodegenerative disorders and cardiovascular disease -
 XX Claim 11; Page 4264-4265; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteoprotective; anticonvulsant; antiallergic; immunosuppressive;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasorelaxant;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihypertensive and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 SQ Sequence 165 AA;
 Query Match 80.6%; Score 141; DB 21; Length 165;
 Best Local Similarity 96.2%; Pred. No. 5, 1e-12;
 Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DCECLGGGRISHOSQDKIHVGYSM 26
 DB 110 DCECLGGGRISHOSQDKIHVGYSM 135
 RESULT 11
 AAB08788
 ID AAB08788 standard; Peptide; 16 AA.
 AC AAB08788;
 XX
 DT 02-JAN-2001 (first entry)
 DE Conserved sequence of mammalian histidine protein phosphatases.
 XX
 KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 KW cell regulation; cell growth; cancer; immune disorder; viral infection;
 KW genetic disorder; heart disease; N-phosphorylation imbalance.
 XX
 OS Mammalia sp.
 XX
 PN WO200052175-A1.
 PD 08-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-EP01774.
 PR 04-MAR-1999; 99DE-1009388.
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Klump S, Kellner R;
 DR WPI; 2000-572187/53.
 XX
 PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
 PT immune disorders, viral infection, genetic disorders, and heart disease
 PT
 PS Claim 2; Page 17; 39pp; English.
 XX

CC The present sequence represents a conserved sequence found in mammalian
 CC histidine phosphatase polypeptides. The human polypeptide has a high
 CC specificity for phosphohistidine and a molecular weight of 13000-15000
 CC Da. The human histidine phosphatase gene is localised at chromosome 9
 CC (9q33). The human histidine phosphatase can be used for diagnosis and
 CC treatment of pathological states of cell regulation and cell growth.
 CC These include cancers, immune disorders, viral infection, genetic
 CC disorders, and heart disease. The histidine phosphatase can also be
 CC used for identifying agonists and antagonists which can be used to
 CC treat conditions associated with N-phosphorylation imbalance.
 XX
 SQ Sequence 16 AA;
 Query Match 52.6%; Score 92; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 3, 1e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DCECLGGGRISHOSQD 16
 DB 1 DCECLGGGRISHOSQD 16
 RESULT 12
 ABB58764
 ID ABB58764 standard; Protein; 135 AA.
 AC ABB58764;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 3084.
 DE
 DE Drosophila melanogaster polypeptide SEQ ID NO 3084.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW
 KW Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02867.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 XX Disclosure; SEQ ID NO 3084; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 135 AA;

XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PR
XX
XX (PEKE) PE CORP NY.
PA
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR
XX N-PSDB; ABL02863.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Disclosure; SEQ ID NO 3072; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 148 AA;

Query Match 37.7%; Score 66; DB 22; Length 148;
Best Local Similarity 38.7%; Pred. No. 0.14;
Matches 12; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHQSODXKIHYVYSMXYGXAOH 33
:|||||:::|:|:|
DB 89 KCLGGGFISNKNEDKWKIKYGCCKTFGEAPH 119

Search completed: May 7, 2003, 19:03:26
Job time : 18.5596 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:00 ; Search time 7.11468 Seconds
(without alignments)
445,900 Million cell updates/sec

Title: US-09-914-831-4

Perfect score: 175
Sequence: 1 DCECLGGCRISHQSDXKIHVYGYSMYGYXAOH 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	52.0	115	2	T21833
2	91	52.0	135	2	A32317
3	90	51.4	149	2	S53876
4	80	45.7	140	2	B32317
5	68	38.9	140	2	S53877
6	51	29.1	147	2	T33531
7	50	28.6	145	2	H86231
8	48.5	27.7	397	2	E53240
9	48	27.4	373	2	G95013
10	48	27.4	391	1	B97887
11	47	26.9	308	1	A34082
12	47	26.9	436	2	D65007
13	47	26.9	436	2	A98032
14	47	26.9	436	2	B85876
15	46.5	26.6	6359	2	T31679
16	46	26.3	428	2	AC0560
17	46	26.3	495	2	B71360
18	46	26.3	822	2	G84681
19	46	26.3	999	2	E64712
20	46	26.3	1001	2	E71806
21	46	26.3	4464	2	D87755
22	45	25.7	275	2	A81284
23	45	25.7	307	2	G90087
24	45	25.7	309	1	XNRCV
25	45	25.7	309	2	H91216
26	45	25.7	309	2	A86063
27	45	25.7	309	2	AD0924
28	45	25.7	394	2	C83840
29	45	25.7	398	2	C82614

30	45	25.7	408	2	B84485	Mutator-like trans
31	44.5	25.4	118	2	A60044	fumarate reductase
32	44.5	25.4	484	2	E82051	pyridine nucleotid
33	44.5	25.4	701	2	T14757	hypothetical prote
34	44	25.1	160	2	F64816	yfia protein - Esc
35	44	25.1	204	2	A10620	probable bacteriop
36	44	25.1	265	2	S34668	modulation protein
37	44	25.1	367	2	T31750	hypothetical prote
38	44	25.1	499	2	C89015	protein B0213.12 [
39	44	25.1	548	2	A54510	63k antigen - nema
40	44	25.1	548	2	A28209	60k filarial anti
41	44	25.1	549	2	T25423	hypothetical prote
42	44	25.1	552	2	T10850	y4hp protein - Rhi
43	44	25.1	557	2	G69958	SNP2 helicase homo
44	44	25.1	599	2	S64136	methylentetracycl
45	44	25.1	733	2	T49679	probable branching

ALIGNMENTS

RESULT 1
T21833
hypothetical protein F36A2.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21833
R:Lenhard, N.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19476
A:Accession: T21833
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-115 <MIL>
A:Cross-references: EMBL:Z81077; PIDN:CAB03070.1; GSPDB:GN00019; CESP:F36A2.8
A:Experimental source: clone F36A2
C:Genetics:
A:Gene: CESP:F36A2.8
A:Map position: 1
A:Introns: 22/3; 52/3

Query Match 52.0%; Score 91; DB 2; Length 115;
Best Local Similarity 54.8%; Pred. No. 1.4e-06;
Matches 17; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
DB 62 KCVGGGRIMHDKIDILVGYGYGRADH 92
QY 3 ECLGGGRISHQSDXKIHVYGYSMYGYXAOH 33
|:||||| | ||||| | | |

RESULT 2
A32317
sex-regulated protein JanA - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 29-Jan-1990 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C:Accession: A32317
R:Yanicoslas, C.; Vincent, A.; Lepesant, J.A.
Mol. Cell. Biol. 9, 2526-2535, 1989
A:Title: Transcriptional and posttranscriptional regulation contributes to the sex-re
A:Reference number: A32317; MUID:89345970; PMID:2503707
A:Accession: A32317
A:Status: preliminary
A:Molecule type: DNA; mRNA
A:Residues: 1-135 <YAN>
A:Cross-references: GB:M27033; NID:q3522914; PIDN:AAC34203.1; PID:q387597
C:Note: the authors translated the codon TTG for residue 98 as Met
C:Genetics:
A:Gene: FlyBase:JanA
A:Cross-references: FlyBase:FBgn0001280

Query Match 52.0%; Score 91; DB 2; Length 135;
Best Local Similarity 51.6%; Pred. No. 1.7e-06;
Matches 16; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSODXKIHVYGYSMYKXAOH 33
 ||||||| : ||||| : |||
 Db 82 ECPGGREHNPEKKYKLYSGFGKADH 112

RESULT 3

sex-regulated protein janus A - fruit fly (Drosophila pseudoobscura)

C:Species: Drosophila pseudoobscura
 C>Date: 14-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Jan-1999
 C:Accession: S53876
 R:Yanicosas, C.; Ferrer, P.; Vincent, A.; Lepesant, J.A.
 M.Ol. Gen. Genet. 246, 549-560, 1995
 A:Title: Separate cis-regulatory sequences control expression of serendipity beta and ja
 A:Reference number: S53876; MUID:95214615; PMID:7700229
 A:Accession: S53876
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-149 <YAN>
 A:Cross-references: GB:S77099; NID:9913977; PID:9913978
 C:Genetics:
 A:Gene: janA
 A:Cross-references: FlyBase:FBgn0015151
 A:Introns: 78/1; 119/3

Query Match 51.4%; Score 90; DB 2; Length 149;
 Best Local Similarity 51.5%; Pred. No. 2.6e-06;
 Matches 17; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSODXKIHVYGYSMYKXAOH 33
 ||||||| : ||||| : |||
 Db 94 DTGCLGGGRIGHNPKKYLKYGSHSTGKADH 126

RESULT 4

sex-regulated protein JanB - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
 C>Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 21-Jul-2000
 C:Accession: B32317
 R:Yanicosas, C.; Vincent, A.; Lepesant, J.A.
 M.Ol. Cell. Biol. 9, 2526-2535, 1989
 A:Title: Transcriptional and posttranscriptional regulation contributes to the sex-regul
 A:Reference number: A32317; MUID:89343970; PMID:2503707
 A:Accession: B32317
 A:Status: preliminary
 A:Molecule type: DNA; mRNA
 A:Residues: 1-140 <YAN>
 A:Cross-references: GB:M27033; NID:93522914; PIDN:AAC34204.1; PID:9387598
 C:Genetics:
 A:Gene: FlyBase:JanB
 A:Cross-references: FlyBase:FBgn0001281

Query Match 45.7%; Score 80; DB 2; Length 140;
 Best Local Similarity 48.3%; Pred. No. 8.3e-05;
 Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 5 LGGGRISHOSODXKIHVYGYSMYKXAOH 33
 ||||||| : ||||| : |||
 Db 90 LGGGRILNEANMKIKITGYSRTGADH 118

RESULT 5

sex-regulated protein janus B - fruit fly (Drosophila pseudoobscura)

C:Species: Drosophila pseudoobscura
 C>Date: 18-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Jan-1999
 C:Accession: S53877
 R:Yanicosas, C.; Ferrer, P.; Vincent, A.; Lepesant, J.A.
 M.Ol. Gen. Genet. 246, 549-560, 1995
 A:Title: Separate cis-regulatory sequences control expression of serendipity beta and ja
 A:Reference number: S53876; MUID:95214615; PMID:7700229

A:Accession: S53877
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-140 <YAN>
 A:Cross-references: GB:S77099; NID:9913977; PID:9913978
 C:Genetics:
 A:Gene: janB
 A:Cross-references: FlyBase:FBgn0014784
 A:Introns: 18/1; 70/1; 111/3

Query Match 38.9%; Score 68; DB 2; Length 140;
 Best Local Similarity 35.5%; Pred. No. 0.0056;
 Matches 11; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSODXKIHVYGYSMYKXAOH 33
 ||||||| : ||||| : |||
 Db 88 KCLGGKMDNDSARKITGHSCKTGAANH 118

RESULT 6

hypothetical protein F58E1.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T33531
 R:Wamsley, P.; Twyman, B.
 A:Description: The sequence of C. elegans cosmid F58E1.
 A:Reference number: 221365
 A:Accession: T33531
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-477 <WAM>
 A:Cross-references: EMBL:AF098995; PIDN:AAC67482.1; GSPDB:GN00020; CESP:F58E1.4
 C:Experimental source: strain Bristol N2; clone F58E1
 C:Genetics:
 A:Gene: CESP:F58E1.4
 A:Map position: 2
 A:Introns: 32/1; 104/1; 269/3; 424/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein T05A6.4

Query Match 29.1%; Score 51; DB 2; Length 477;
 Best Local Similarity 47.4%; Pred. No. 7.9;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSODXKI 19
 ||||| : ||| : |||
 Db 170 DCECGSGRVITGYLRDAKV 188

RESULT 7

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86231
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Con, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Hultar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H86231
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-454 <STO>
 A:Cross-references: GB:AE005172; NID:92160169; PIDN:AAB60732.1; GSPDB:GN00141
 C:Genetics:

A:Map position: 1

Query Match 28.6%; Score 50; DB 2; Length 454;

Best Local Similarity 35.7%; Pred. No. 11;

Matches 10; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 4 C1GCGRISHOSQDKIHVGYSMXYGXA 31

DB 179 CCRGGVLTSMQSDAHNVSAFHTVGS 206

RESULT 8

E53240

C:Species: Ambrosia artemisiifolia (common ragweed)

C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Nov-2000

C:Accession: E53240; A46469

R:Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.

A:Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia

A:Reference number: A53240; MUID:92234570; PMID:1809687

A:Accession: E53240

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-397 <GRD>

A:Cross-references: GB:M80561

R:Rogers, B.L.; Morgenstern, J.P.; Griffith, I.J.; Yu, X.B.; Counsell, C.M.; Brauer, A.W.

J. Immunol. 147, 2547-2552, 1991

A:Title: Complete sequence of the allergen Amb alpha II. Recombinant expression and reac

A:Reference number: A46469; MUID:92013060; PMID:1717566

A:Accession: A46469

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-69, 'N', '1-397 <ROG>

A:Cross-references: GB:M80561; NID:9166446; PIDN:AAA32671.1; PID:9166447

A:Note: sequence extracted from NCBI backbone (NCBIN:58755, NCBIP:58758)

C:Superfamily: peccate lyase LAM59

C:Keywords: pollen

Query Match 27.7%; Score 48.5; DB 2; Length 397;

Best Local Similarity 41.2%; Pred. No. 16;

Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

OY 1 DCEGCGRIS-----HOSQDKIHVGYGS 25

DB 171 DYRVLPGRKIKSNGCPAIRHOSDGDALHVTGS 204

RESULT 9

G95013

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C:Accession: G95013

R:Retzlaff, H.; Neilsen, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: G95013

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-373 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK74304.1; PID:914971585; GSPDB:GN00164; TIGR:SP4

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0118

C:Superfamily: probable membrane protein YDL033c

Query Match 27.4%; Score 48; DB 2; Length 373;

Best Local Similarity 45.0%; Pred. No. 18;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSQDKIHVY 22

DB 353 ECLGGGLIDNAYRDCVCQY 372

RESULT 10

B97887

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001

C:Accession: B97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: B97887

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK98926.1; PID:915457661; GSPDB:GN00174

C:Genetics:

A:Gene: trmD

C:Superfamily: probable membrane protein YDL033c

C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 27.4%; Score 48; DB 2; Length 391;

Best Local Similarity 45.0%; Pred. No. 19;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSQDKIHVY 22

DB 371 ECLGGGLIDNAYRDCVCQY 390

RESULT 11

A34082

C:Species: Salmonella typhimurium

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: A34082; A05077

R:Feld, M.J.; Nguyen, D.C.; Armstrong, F.B.

Biochemistry 28, 5306-5310, 1989

A:Title: Amino acid sequence of Salmonella typhimurium branched-chain amino acid amin

A:Reference number: A34082; MUID:89352621; PMID:2669973

A:Accession: A34082

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-308 <FRD>

A:Function:

A:Description: catalyzes the reversible transamination of alpha-amino groups from leu

A:Pathway: branched-chain amino acid biosynthesis

C:Superfamily: branched-chain-amino-acid aminotransferase

C:Keywords: aminotransferase; branched-chain amino acid biosynthesis; phosphoprotein;

F;159/Binding site: pyridoxal phosphate (Lys) (covalent) *Status Predicted

Query Match 26.9%; Score 47; DB 1; Length 308;

Best Local Similarity 31.6%; Pred. No. 21;

Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

OY 13 QSQDKRISHVGYSMXYGXA 31

DB 16 RNEDAKVHMSALHTGTS 34

RESULT 12

acetyl-CoA C-acyltransferase (EC 2.3.1.16) - Escherichia coli (strain K-12)
 D65007
 C:Species: Escherichia coli
 C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: D65007
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: D65007
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-436 <BLAT>
 A:Cross-references: GB:AE000322; GB:U00096; NID:g1788672; PID:AC75402.1; PID:g1788683;
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: long-chain-fatty acid beta-oxidation multienzyme complex beta chain
 C:Keywords: acyltransferase; coenzyme A

Query Match 26.9%; Score 47; DB 2; Length 436;
 Best Local Similarity 37.0%; Pred. No. 30;
 Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHQSODXKIHVGYSMXYG 29
 DB 365 EALGRAHATGEVDSKFNVLGGSIATYG 391

RESULT 13

probable acyltransferase [imported] - Escherichia coli (strain O157:H7, substrain R1MD
 A98032
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
 C:Accession: A98032
 R:Hayashi, T.; Makino, K.; Ohishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: A98032
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA36648.1; PID:g13362695; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain R1MD 0509952
 C:Genetics:
 A:Gene: EGS3225
 C:Superfamily: long-chain-fatty acid beta-oxidation multienzyme complex beta chain

Query Match 26.9%; Score 47; DB 2; Length 436;
 Best Local Similarity 37.0%; Pred. No. 30;
 Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHQSODXKIHVGYSMXYG 29
 DB 365 EALGRAHATGEVDSKFNVLGGSIATYG 391

RESULT 14

probable acyltransferase 23605 [imported] - Escherichia coli (strain O157:H7, substrain
 B85876
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
 C:Accession: B85876
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85876

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <STO>

A:Cross-references: GB:AE005174; NID:g12516706; PIDN:ACG57470.1; GSPDB:GN00145; UMGCP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: 23605

C:Superfamily: long-chain-fatty acid beta-oxidation multienzyme complex beta chain

Query Match 26.9%; Score 47; DB 2; Length 436;
 Best Local Similarity 37.0%; Pred. No. 30;
 Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 3 ECLGGGRISHQSODXKIHVGYSMXYG 29
 DB 365 EALGRAHATGEVDSKFNVLGGSIATYG 391

RESULT 15

bacitracin synthetase 3 - Bacillus licheniformis
 T31679
 C:Species: Bacillus licheniformis
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
 C:Accession: T31679
 R:Kron, D.; Klens, A.; Schorgendorfer, K.; Marahiel, M.A.
 Chem. Biol. 4, 927-937, 1997
 A>Title: The bacitracin biosynthesis operon of Bacillus licheniformis ATCC 10716: mol
 A:Reference number: Z21058; MUID:98089193; PMID:9427658
 A:Accession: T31679
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-6359 <KON>
 A:Cross-references: EMBL:AF007865; NID:g4464275; PID:g2982196; PIDN:AC06348.1
 C:Genetics:
 A:Gene: bacC
 C:Superfamily: acetate-CoA ligase homology; acyl carrier protein homology
 C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
 F:505-946/Domain: acetate-CoA ligase homology <ACLI>
 F:964-1032/Domain: acyl carrier protein homology <ACLP1>
 F:1542-1978/Domain: acetate-CoA ligase homology <ACLI2>
 F:11996-2063/Domain: acyl carrier protein homology <ACRP>
 F:3046-3483/Domain: acetate-CoA ligase homology <ACLI3>
 F:3500-3568/Domain: acyl carrier protein homology <ACRP3>
 F:4078-4526/Domain: acetate-CoA ligase homology <ACLI4>
 F:4542-4609/Domain: acyl carrier protein homology <ACRP4>
 F:5593-6032/Domain: acetate-CoA ligase homology <ACLI5>
 F:6050-6118/Domain: acyl carrier protein homology <ACRP5>
 F:3532/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 26.6%; Score 46.5; DB 2; Length 6359;
 Best Local Similarity 39.3%; Pred. No. 5.7e+02;
 Matches 11; Conservative 3; Mismatches 5; Indels 9; Gaps 1;

QY 5 LGGGR-----ISHQSODXKIHVYG 23
 DB 719 LGGGRASIPVRRVILNHVGRDLIHVYG 746

Search completed: May 7, 2003, 19:07:33
 Job time : 9.11468 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 18:54:40 ; Search time 3.93578 Seconds
(without alignments)
347.763 Million cell updates/sec

Title: US-09-914-831-4

Perfect score: 175
Sequence: 1 DCECGGGRISHQSDXKIHVGYGMYGXQAH 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	57.7	146	1	JANA_DROPS
2	91	52.0	135	1	P20348 drosophila
3	80	45.7	140	1	JANA_DROME
4	70	40.0	140	1	JANA_DROPS
5	49.5	28.3	1856	1	MCA_HUMAN
6	48.5	27.7	397	1	MPA2_AMBAR
7	47	26.9	436	1	YFCY_ECOLI
8	46.5	26.6	6359	1	BACC_BACLI
9	46	26.3	485	1	Y136_TREPA
10	46	26.3	890	1	KPCN_HUMAN
11	46	26.3	999	1	MPD_HELPY
12	46	26.3	1001	1	MPD_HELPY
13	46	26.3	4568	1	DYHC_CAEEL
14	45	25.7	308	1	ILVE_ECOLI
15	45	25.7	308	1	ILVE_ECOLI
16	45	25.7	972	1	HGBA_HABED
17	45	25.7	2282	1	ZAN_RABIT
18	44	25.1	160	1	YBIA_ECOLI
19	44	25.1	343	1	HXDA_BRARE
20	44	25.1	548	1	SYN_BRUMA
21	44	25.1	552	1	Y4HP_RHISN
22	44	25.1	557	1	YOHM_BACSU
23	44	25.1	599	1	MTHS_YEAST
24	44	25.1	747	1	ORPB_HUMAN
25	44	25.1	1272	1	UBP2_YEAST
26	43.5	24.9	389	1	WNIB_HUMAN
27	43.5	24.9	389	1	WNIB_MOUSE
28	43.5	24.9	433	1	UROK_PAPCY
29	43.5	24.9	5255	1	BACA_BACLI
30	43	24.6	230	1	YWOI_CAEEL
31	43	24.6	382	1	SUBT_BACAM
32	43	24.6	446	1	TBBI_TRIYI
33	43	24.6	559	1	PACT_MOUSE

34	43	24.6	584	1	EXON_HCVNA	P16789 human cytom
35	43	24.6	623	1	PMT1_DROME	P51022 drosophila
36	43	24.6	908	1	DM3A_MOUSE	O88508 mus musculus
37	43	24.6	909	1	DM3A_HUMAN	O956K1 homo sapien
38	43	24.6	1229	1	C1BB_BACTU	O45733 bacillus th
39	43	24.6	1231	1	C1BD_BACTU	O92425 bacillus th
40	43	24.6	1279	1	BCSC_PSEPL	P58937 pseudomonas
41	43	24.6	1476	1	A2M1_MOUSE	P28665 mus musculus
42	42.5	24.3	61	1	SIX2_BURST	P82812 buthus sind
43	42.5	24.3	114	1	FRDD_HAEIN	P44891 haemophilus
44	42.5	24.3	505	1	AMID_PSECL	P27765 pseudomonas
45	42.5	24.3	804	1	RIR1_HSV6U	P52343 human herpe

ALIGNMENTS

RESULT 1
JANA_DROPS STANDARD: PRT; 146 AA.
AC P54364;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sex-regulated protein janus-A.
GN JANA.
OS Drosophila pseudoobscura (Fruit fly).
OC Insecta; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95214615; PubMed=7700229;
RA Yancostas C., Ferrer P., Vincent A., Lepesant J.-A.;
RT beta and janus A, two immediately adjacent Drosophila genes.*;
RL Mol. Gen. Genet. 246:549-560(1995).
CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-B.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL; S77099; AB33912.2; -;
DR FlyBase; FBgn001515L; DpseJana.
SQ SEQUENCE 146 AA; 16291 MW; 02F7DEC570070D7B CRC64;
SO
Query Match 57.7%; Score 101; DB 1; Length 146;
Best Local Similarity 54.5%; Pred. No. 3.1e-08;
Matches 18; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
OY 1 DCECGGGRISHQSDXKIHVGYGMYGXQAH 33
DB 91 DTECGGGRISHQSDXKIHVGYGMYGXQAH 123
RESULT 2
JANA_DROME STANDARD: PRT; 135 AA.
AC P20348; G9VAB6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sex-regulated protein janus-A.
GN JANA OR CG7933.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha: Ephydroidea: Drosophilidae: Drosophila.
 NCBI_TaxID=7227;
 RN SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=89343970; PubMed=2503707;
 RA Yancoskas C., Vincent A., Lepesant J.-A.;
 RT "Transcriptional and posttranscriptional regulation contributes to
 the sex-regulated expression of two sequence-related genes at the
 Janus locus of *Drosophila melanogaster*.";
 RL Mol. Cell. Biol. 9:2526-2535(1989).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jitali M., Kalush F., Kapen G.H., Ke Z., Kenison J.A., Kelchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- TISSUE SPECIFICITY: SOMATIC AND GERM LINE CELLS.
 CC -1- DEVELOPMENTAL STAGE: THE NON-SEX-SPECIFIC TRANSCRIPT IS PRESENT AT
 ALL STAGES.
 CC -1- MISCELLANEOUS: TRANSCRIPTION OF JANUS GIVES RISE TO TWO
 CC SEX-SPECIFIC AND TO ONE NON-SEX-SPECIFIC TRANSCRIPTS.
 CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-B.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M27033; AAC34203.1; -
 CC EMBL: M27033; AAC34202.1; ATL-INT.
 CC EMBL: AE003772; AAF56997.1; -
 CC PIR: A32317; A32317.
 CC FLYBase: FBgn0001280; jana:

SQ SEQUENCE 135 AA; 15220 MW; 2720237CE7F7F3132 CRC64;
 Query Match 52.0%; Score 91; DB 1; Length 135;
 Best Local Similarity 51.6%; Pred. No. 9.1e-07;
 Matches 16; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
 QY 3 ECLGGRISHSQDQKKIHVGYSMXYGXAQH 33
 DB 82 ECPGGRIENHEKRYLVYGVSGGFGKADH 112
 RESULT 3
 JANUS-DROME
 ID JANUS-DROME STANDARD; PRT; 140 AA.
 AC P20349; G9VAB7;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sex-regulated protein janus-B.
 GN JANUS OR CG7931.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=89343970; PubMed=2503707;
 RA Yancoskas C., Vincent A., Lepesant J.-A.;
 RT "Transcriptional and posttranscriptional regulation contributes to
 the sex-regulated expression of two sequence-related genes at the
 Janus locus of *Drosophila melanogaster*.";
 RL Mol. Cell. Biol. 9:2526-2535(1989).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jitali M., Kalush F., Kapen G.H., Ke Z., Kenison J.A., Kelchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- TISSUE SPECIFICITY: GERM LINE CELLS OF ADULT MALES.
 CC -1- DEVELOPMENTAL STAGE: FROM THE THIRD-INSTAR LARVAL STAGE TO THE
 CC ADULT STAGE.
 CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-A.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M27033; AAC34204.1; .
 DR EMBL: AE003772; AAF5696.1; .
 DR PIR: B32317; B32317
 DR Flybase: FBgn0001281; JanB.
 SQ SEQUENCE 140 AA; 15859 MW; 5750CC28BB61D7FE CRC64;
 Query Match 45.7%; Score 80; DB 1; Length 140;
 Best Local Similarity 48.3%; Pred. No. 4.2e-05;
 Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
 QY 5 LGGGRISHOSQDKIHVGYSMXYGAQH 33
 DB 90 LGGGRILNEAKNKIKYGTSTRTFGADH 118
 ID JANB_DROPS STANDARD: PRT: 140 AA.
 AC P54365;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sex-regulated protein Janus-B.
 GN JANB.
 OS *Drosophila pseudoobscura* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_Taxid=7237;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95214615; PubMed=7700229;
 RA Yanicostas C., Ferrer P., Vincent A., Lepesant J.-A.;
 RT "Separate cis-regulatory sequences control expression of serendipity
 RL beta and janus A, two immediately adjacent *Drosophila* genes.";
 RL Mol. Gen. Genet. 246:549-560(1995).
 CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-A.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S77099; AAB3913.2; .
 DR Flybase: FBgn0014784; DPse\JanB.
 SQ SEQUENCE 140 AA; 15812 MW; 01B87FE1F0A5F53A CRC64;
 Query Match 40.0%; Score 70; DB 1; Length 140;
 Best Local Similarity 35.5%; Pred. No. 0.0014;
 Matches 11; Conservative 7; Mismatches 13; Indels 0; Gaps 0;
 QY 3 ECLGGGRISHOSQDKIHVGYSMXYGAQH 33
 DB 88 KCLGGKMDNDESAKRIHGSCKTFGAANH 118

RESULT 5
 MGA_HUMAN STANDARD: PRT: 1856 AA.
 ID MGA_HUMAN
 AC Q43451;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Maltase-glucoamylase, intestinal [includes: Maltase (EC 3.2.1.20)
 DE (Alpha-glucosidase); glucoamylase (EC 3.2.1.3) (Glucoan 1,4-alpha-
 DE glucosidase)].
 GN MGA OR MGA OR MGA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Small intestine;
 RX MEDLINE=98112863; PubMed=9446624;
 RA Nichols B.L., Eldering J.A., Avery S.E., Hahn D., Quaroni A.,
 RA Sterchi E.E.;
 RT "Human small intestinal maltase-glucoamylase cDNA cloning. Homology
 RT to sucrase-isomaltase.";
 RL J. Biol. Chem. 273:3076-3081(1998).
 RN [2]
 RP REVISIONS TO 776: 1049; 1100; 1541; 1612 AND 1811.
 RA Nichols B.L., Eldering J.A., Avery S.E., Hahn D., Quaroni A.,
 RA Sterchi E.E.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP CHARACTERIZATION.
 RC TISSUE=Small intestine; mucosa;
 RX MEDLINE=89066802; PubMed=3143729;
 RA Naim H.Y., Sterchi E.E., Lentze M.J.;
 RT "Structure, biosynthesis, and glycosylation of human small intestinal
 RL maltase-glucoamylase.";
 RL J. Biol. Chem. 263:19709-19717(1988).
 RN [4]
 RP SUFFUTION.
 RX MEDLINE=88082658; PubMed=3121301;
 RA Danielson E.M.;
 RT "Tyrosine sulfation, a post-translational modification of microvillar
 RT enzymes in the small intestinal enterocyte.";
 RL EMBO J. 6:2891-2896(1987).
 CC -1- FUNCTION: MAY SERVE AS AN ALTERNATE PATHWAY FOR STARCH DIGESTION
 CC WHEN LUMINAL ALPHA-AMYLASE ACTIVITY IS REDUCED BECAUSE OF
 CC IMMATUREITY OR MALNUTRITION. MAY PLAY A UNIQUE ROLE IN THE
 CC DIGESTION OF MALTED DIETARY OLIGOSACCHARIDES USED IN FOOD
 CC MANUFACTURING.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
 CC linked D-glucose residues with release of D-glucose.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. BRUSH BORDER.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SMALL INTESTINE, GRANULOCYTE, AND
 CC KIDNEY BUT NOT IN SALIVARY GLAND OR PANCREAS.
 CC -1- PTM: N- AND O-GLYCOSYLATED.
 CC -1- PTM: DOES NOT UNDERGO INTRACELLULAR OR EXTRACELLULAR PROTEOLYTIC
 CC CLEAVAGE.
 CC -1- PTM: SULFATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 31 OF GLYCOSYL HYDROLASES.
 CC -1- SIMILARITY: CONTAINS 2 P-TYPE (TREPOLI) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-97349980; PubMed-9205837;
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
 RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horikuchi T.;
 RT "Construction of a contiguous 874-Kb sequence of the Escherichia coli
 K-12 genome corresponding to 50.0-68.8 min on the linkage map and
 analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + acetyl-CoA = CoA + 3-oxoacyl-CoA.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.
 CC -1- CAUTION: Ref 2 sequence differs from that shown due to frameshifts
 in positions 12 and 303.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000322; AAC75402.1; -
 DR EMBL: D90864; BAA16196.1; ALT_FRAME.
 DR EMBL: D90864; BAA16197.1; ALT_FRAME.
 DR EMBL: D90865; BAA16202.1; ALT_FRAME.
 DR EMBL: D90865; BAA16203.1; ALT_FRAME.
 DR HSP: P27796; IAFY.
 DR Ecocore: EG14128; yfey.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam: PF02803; thiolase; 1.
 DR PROSITE: PS00098; THIOLASE_1; 1.
 DR PROSITE: PS0037; THIOLASE_2; 1.
 DR PROSITE: PS00099; THIOLASE_3; 1.
 KW Hypothetical protein; Transferase; Acyltransferase; Complete proteome.
 FT ACT_SITE 99 99 SUBSTRATE BINDING (BY SIMILARITY).
 FT ACT_SITE 422 422 BASE (BY SIMILARITY).
 SQ SEQUENCE 436 AA: 46530 MW: 586983172C16299 CRC64;
 QY
 Db 3 ECLGGRISSHODKIHYYGSKMYG 29
 365 EALGRAHATGEVDSKFNVLGGSIAYG 391
 Query Match 26.9%; Score 47; DB 1; Length 436;
 Best Local Similarity 37.0%; Pred. No. 12;
 Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

DR 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Bactiracin synthetase 3 (BA3) [Includes: ATP-dependent isoleucine
 DE adenylation (IleA) (Isoleucine activase); ATP-dependent D-phenylalanine
 DE adenylation (D-PheA) (D-phenylalanine activase); ATP-dependent histidine
 DE adenylation (HisA) (Histidine activase); ATP-dependent D-aspartate
 DE adenylation (D-AspA) (D-aspartate activase); ATP-dependent asparagine
 DE adenylation (AsnA) (Asparagine activase); Aspartate racemase
 DE (EC 5.1.1.13); Phenylalanine racemase (ATP hydrolyzing)
 DE (EC 5.1.1.11)].
 GN BAC.
 OS Bacillus licheniformis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1402;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 10716;
 RX MEDLINE-96089193; PubMed-9427658;
 RA Konz D., Klenz A., Schoergerdorfer K., Marahiel M.A.;
 RT "The Bactiracin biosynthesis operon of Bacillus licheniformis ATCC
 10716: molecular characterization of three multi-modular peptide
 synthetases.";
 RL Chem. Biol. 4:927-937(1997).
 CC -1- FUNCTION: INDUCES PEPTIDE SYNTHESIS. ACTIVATES AND INCORPORATES
 FIVE AMINO ACIDS, FORMS A THIAZOLINE RING BETWEEN THE FIRST TWO
 AMINO ACIDS AND INCORPORATES A D-GLUTAMINE IN THE FOURTH POSITION.
 CC -1- CATALYTIC ACTIVITY: L-aspartate = D-aspartate.
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine = AMP + diphosphate + D-
 phenylalanine.
 CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTHEINES
 (POTENTIAL).
 CC -1- PATHWAY: Cyclic peptide antibiotic bactiracin biosynthesis.
 CC -1- SUBUNIT: LARGE MULTISUBUNIT COMPLEX OF BA1, BA2 AND BA3.
 CC -1- DOMAIN: CONSISTS OF FIVE MODULES WITH TWO EPIMERIZATION DOMAINS IN
 THE SECOND AND FOURTH MODULES, AND A PUTATIVE C-TERMINAL
 THIOESTERASE DOMAIN. EACH MODULE INCORPORATES ONE AMINO ACID INTO
 THE PEPTIDE PRODUCT AND CAN BE FURTHER SUBDIVIDED INTO DOMAINS
 RESPONSIBLE FOR SUBSTRATE ADENYLATION, THIOLATION, CONDENSATION
 (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND
 N METHYLATION (OPTIONAL).
 CC -1- MISCELLANEOUS: BACTIRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC
 DODECAPETIDES, THAT DIFFER BY ONE OR TWO AMINO ACIDS. THE MOST
 ABUNDANT IS BACTIRACIN A, A BRANCHED CYCLIC DODECAPETIDE. IT
 CONTAINS AN N-TERMINAL LINEAR PENTAPEPTIDE MOIETY (ILE-CYS-LEU-D-
 GLU-ILE) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION
 PRODUCT AND A C-TERMINAL HEPTAPEPTIDE RING (LYS-D-ORN-ILE-D-PHE-
 HIS-D-ASP-ASN), IN WHICH THE FREE ALPHA-CARBOXY GROUP OF THE C-
 TERMINAL ASN IS BOUND TO THE EPSILON-AMINO GROUP OF LYSINE. IT
 CONTAINS FOUR AMINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7,
 PHE-9, AND ASP-11).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF007865; AAC06348.1; -
 DR HSP: P14687; IAMU.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatn.
 DR InterPro: IPR003880; Pantine attach.
 DR InterPro: IPR001031; thioesterase.
 DR Pfam: PF00501; AMP-binding; 5.
 DR Pfam: PF00550; PP-binding; 5.
 DR Pfam: PF00668; Condensation; 7.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PRINTS: PR00154; AMPBINDING.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 4.

DR PROSITE: PS00455; AMP_BINDING: 5.
 DR PROSITE: PS50075; ACP_DOMAIN: 5.
 KM Ligase; Isomerase; Hydrolase; Antibiotic biosynthesis;
 KM Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT REPEAT 461 1034 DOMAIN 1 (ISOLEUCINE-ACTIVATING).
 FT REPEAT 1517 2064 DOMAIN 2 (D-PHENYLLAMINE-ACTIVATING).
 FT REPEAT 2999 3570 DOMAIN 3 (HISTIDINE-ACTIVATING).
 FT REPEAT 4047 4612 DOMAIN 4 (D-ASPARTIC ACID-ACTIVATING).
 FT REPEAT 5549 6129 DOMAIN 5 (ASPARAGINE-ACTIVATING).
 FT DOMAIN 966 1033 ACTL CARRIER (ACP) 1.
 FT DOMAIN 1998 2064 ACTL CARRIER (ACP) 2.
 FT DOMAIN 3502 3569 ACTL CARRIER (ACP) 3.
 FT DOMAIN 4544 4610 ACTL CARRIER (ACP) 4.
 FT DOMAIN 6052 6119 ACTL CARRIER (ACP) 5.
 FT BINDING 996 996 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 2028 2028 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 3532 3532 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 4574 4574 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 6082 6082 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 SQ SEQUENCE 6359 AA; 722923 MW; 82A273C546253074 CRC64;

Query Match 26.6%; Score 46.5; DB 1; Length 6359;
 Best Local Similarity 39.3%; Pred. No. 2.3e+02;
 Matches 11; Conservative 3; Mismatches 5; Indels 9; Gaps 1;

OY 5 LGGR-----ISHQSDXKIHVYG 23
 ||| | :| | |||||
 DB 719 LGGRASIPHYKVLNVGDKLIHYVG 746

RESULT 9
 Y136 TREPA STANDARD; PRT; 485 AA.
 ID 083172;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Hypothetical lipoprotein TP0136 precursor.
 GN TP0136.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Kralak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 spirochete.";
 RL Science 281:375-388(1998).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Potential).
 CC -1- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB001199; AAC65137.1; ALT_INIT.
 DR TIGR: TP0136;
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;
 FT Complete proteome. 1 23 POTENTIAL.

FT CHAIN 24 485 HYPOTHETICAL LIPOPROTEIN TP0136.
 FT LIPID 24 24 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT DOMAIN 164 178 GLY/SER-RICH.
 FT DOMAIN 196 210 GLY/SER-RICH.
 FT DOMAIN 253 267 GLY/SER-RICH.
 FT DOMAIN 318 327 POLY-SER.
 FT DOMAIN 444 447 POLY-SER.
 SQ SEQUENCE 485 AA; 48984 MW; C7ACCEDC7DC5CED CRC64;

Query Match 26.3%; Score 46; DB 1; Length 485;
 Best Local Similarity 40.9%; Pred. No. 20;
 Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 1 DCECLGGRISHQSDXKIHVY 22
 :| | | | | :| | | | |
 DB 162 NCGCGGGGSSSSSSSCIHV 183

RESULT 10
 KPCN_HUMAN STANDARD; PRT; 890 AA.
 ID KPCN_HUMAN
 AC 094806;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein kinase C, nu type (EC 2.7.1.1-) (PKC-nu) (Protein kinase
 DE EPK2).
 GN KPCN OR EPK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=embryo;
 RX MEDLINE=99249819; PubMed=10231560;
 RA Hayashi A., Seki N., Hattori A., Kozuma S., Salto T.;
 RT "PKCnu, a new member of the protein kinase C family, composes a fourth
 RT subfamily with PKCmu.";
 RT Biochim. Biophys. Acta 1450:99-106(1999).
 RL Biochim. Biophys. Acta 1450:99-106(1999).
 CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT,
 CC SERINE- AND THREONINE-SPECIFIC ENZYME. PKC IS ACTIVATED BY
 CC DIACYLGLYCEROL WHICH IN TURN PHOSPHORYLATES A RANGE OF CELLULAR
 CC PROTEINS. PKC ALSO SERVES AS THE RECEPTOR FOR PHORBOL ESTERS, A
 CC CLASS OF TUMOR PROMOTERS.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB015982; BAA36514.1;
 DR HSSP: P28867; IPTQ.
 DR Genew: HGNC:9408; PKCN.
 DR MIM: 607077;
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR pfam: PF00069; pkinase; 1.
 DR pfam: PF00130; DAG_PE-bind; 2.
 DR pfam: PF00169; PH; 1.
 DR PRINTS: PR00008; DAGPEDOMAIN.
 DR Prodom: PD000001; Euk_pkinase; 1.

DR SMART: SM00109; CL: 2.
 DR SMART: SM00233; PH: 1.
 DR SMART: SM00220; S_TKC: 1.
 DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE: PS00081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase; Serine/threonine-protein kinase;
 FM Phorbol-ester binding; Zinc; Repeat.
 FT DOMAIN 155 204 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 272 321 PHORBOL-ESTER AND DAG BINDING 2.
 FT DOMAIN 416 532 PH.
 FT DOMAIN 576 832 PROTEIN KINASE.
 FT NP_BIND 582 590 ATP (BY SIMILARITY).
 FT BINDING 605 605 ATP (BY SIMILARITY).
 FT ACT SITE 699 699 BY SIMILARITY.
 SQ SEQUENCE 890 AA; 100470 MW; 66D5E7E7235064F5 CRC64;

Query Match 26.3%; Score 46; DB 1; Length 890;
 Best Local Similarity 34.6%; Pred. No. 37;
 Matches 9; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

OY 5 LGGGRISHSQDXXK--IHVYGYSMY 28
 DB 848 IGRVYTHESDARWEIHAYTHMLVY 873

RESULT 11
 MFD_HELPY
 ID MFD_HELPY STANDARD: PRT: 999 AA.
 AC 026066;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription-repair coupling factor (TRCF).
 GN MFD OR HP1541.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khakh H.G., Glodek A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Uitterlinden T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 pylori.";
 RL Nature 388:539-547(1997).
 CC -I- FUNCTION: NECESSARY FOR STRAND-SPECIFIC REPAIR. A LESION IN THE
 CC TEMPLATE STRAND BLOCKS THE RNA POLYMERASE COMPLEX (RNAP). THE
 CC RNAP-DNA-RNA COMPLEX IS SPECIFICALLY RECOGNIZED BY TRCF WHICH
 CC RELEASES RNAP AND THE TRUNCATED TRANSCRIPT. THE TRCF MAY REPLACE
 CC RNAP AT THE LESION SITE AND THEN RECRUIT THE UVRA/B/C REPAIR
 CC SYSTEM (BY SIMILARITY).
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE UVB FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HELICASE
 CC FAMILY. RECG SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000652; AAD08581.1; -.
 DR TIGR: HP1541; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR004576; Mfd.
 DR InterPro: IPR005118; TRCF.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR Pfam: PF03461; TRCF; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC_C; 1.
 DR TIGRfams: TIGR00580; mfd; 1.
 KW Helicase; DNA repair; ATP-binding; DNA-binding; Complete proteome.
 FT NP_BIND 512 519 ATP (POTENTIAL).
 FT SITE 609 612 DEAD BOX.
 SQ SEQUENCE 999 AA; 113090 MW; EE550EFB2A2946E5 CRC64;

Query Match 26.3%; Score 46; DB 1; Length 999;
 Best Local Similarity 38.5%; Pred. No. 41;
 Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY 1 DCECLGGRISHSQDXXKIHVYGYSM 26
 DB 829 DLEIRGGGMLGQDQSGHIKNIGVAL 854

RESULT 12
 MFD_HELPY
 ID MFD_HELPY STANDARD: PRT: 1001 AA.
 AC 092057;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription-repair coupling factor (TRCF).
 GN MFD OR JHP1458.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=85963;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99120557; PubMed=9923682;
 RX Alm R.A., Ling L.-S.L., Molr D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guld B.C., deJonge B.L., Carmel G.,
 RA Tumilino P.J., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC -I- FUNCTION: NECESSARY FOR STRAND-SPECIFIC REPAIR. A LESION IN THE
 CC TEMPLATE STRAND BLOCKS THE RNA POLYMERASE COMPLEX (RNAP). THE
 CC RNAP-DNA-RNA COMPLEX IS SPECIFICALLY RECOGNIZED BY TRCF WHICH
 CC RELEASES RNAP AND THE TRUNCATED TRANSCRIPT. THE TRCF MAY REPLACE
 CC RNAP AT THE LESION SITE AND THEN RECRUIT THE UVRA/B/C REPAIR
 CC SYSTEM (BY SIMILARITY).
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE UVB FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HELICASE
 CC FAMILY. RECG SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001567; AAD07030.1; -.
 DR InterPro: IPR001410; DEAD.

DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR004576; Mfd.
 DR InterPro: IPR005118; TRCF.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR Pfam: PF03461; TRCF; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR TIGRfam: TIGR00580; mfd; 1.
 KW Helicase; DNA repair; ATP-binding; DNA-binding; Complete proteome.
 FT NP_BIND 512 519 ATP (POTENTIAL).
 FT SITE 611 614 DEEH BOX.
 SQ SEQUENCE 1001 AA; 113436 MW; D5192A7AF29631A CRC64;
 Query Match 26.3%; Score 46; DB 1; Length 1001;
 Best Local Similarity 38.5%; Pred. No. 41;
 Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY 1 DCELGGRISHOSQDKIHVGYGSM 26
 Db 831 DLEIRGGNLLGDDQSGHINICVAL 856

RESULT 13
 DYHC_CAEEL STANDARD; PRT; 4568 AA.
 ID DYHC_CAEEL
 AC 019020;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dynein heavy chain, cytosolic (DYHC).
 GN DHC-1 OR T21E12.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=96114101; PubMed=8674131;
 RA Lye R.J., Wilson R.K., Waterston R.H.;
 RT "Genomic structure of a cytoplasmic dynein heavy chain gene from the
 RT nematode *Caenorhabditis elegans*.";
 RL Cell Motil. Cytoskeleton 32:26-36(1995).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Du Z., Magyl L.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
 CC ORGANELLES ALONG MICROTUBULES.
 CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
 CC INTERMEDIATE AND LIGHT CHAINS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L33360; AAC37251.1; -
 CC DR EMBL: U80440; AAK21472.1; -
 CC DR Wormpep; T21E12.4; CE23997.
 DR InterPro: IPR003593; AAA_ATPase.

DR InterPro: IPR004273; Dynein_heavy.
 DR Pfam: PF03028; Dynein_heavy; 1.
 DR SMART: SM00382; AAA; 1.
 KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
 FT DOMAIN 587 652 COILED COIL (POTENTIAL).
 FT DOMAIN 814 844 COILED COIL (POTENTIAL).
 FT DOMAIN 1241 1274 COILED COIL (POTENTIAL).
 FT DOMAIN 1324 1340 COILED COIL (POTENTIAL).
 FT DOMAIN 1559 1591 COILED COIL (POTENTIAL).
 FT DOMAIN 1964 1992 COILED COIL (POTENTIAL).
 FT DOMAIN 3132 3229 COILED COIL (POTENTIAL).
 FT DOMAIN 3339 3432 COILED COIL (POTENTIAL).
 FT DOMAIN 3707 3739 COILED COIL (POTENTIAL).
 FT DOMAIN 4359 4386 COILED COIL (POTENTIAL).
 FT NP_BIND 1865 1872 ATP (POTENTIAL).
 FT NP_BIND 2163 2170 ATP (POTENTIAL).
 FT NP_BIND 2537 2544 ATP (POTENTIAL).
 FT NP_BIND 2880 2887 ATP (POTENTIAL).
 SQ SEQUENCE 4568 AA; 521568 MW; 028E52684F381676 CRC64;
 Query Match 26.3%; Score 46; DB 1; Length 4568;
 Best Local Similarity 36.0%; Pred. No. 2e+02;
 Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 8 GRISHOSQDKIHVGYGSMYXGAQ 32
 Db 1056 GRUVFDTQTRKRVFPVSVEYGAQ 1080

RESULT 14
 ILVE_ECOLI STANDARD; PRT; 308 AA.
 ID ILVE_ECOLI
 AC P00510; Q47299;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Branched-chain amino acid aminotransferase (EC 2.6.1.42) (Transaminase
 DE B) (BCAT).
 GN ILVE OR B3770 OR Z5281 OR ECS4704.
 OS Escherichia coli.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 ON NCBI_TaxID=562, 83334;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=85289113; PubMed=3897211;
 RA Kuramitsu S., Ogawa T., Ogawa H., Kagamiyama H.;
 RT "Branched-chain amino acid aminotransferase of *Escherichia coli*:
 RT nucleotide sequence of the *ilvE* gene and the deduced amino acid
 RT sequence.";
 RL J. Biochem. 97:993-999(1985).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=87174741; PubMed=3550695;
 RA Lawther R.P., Wek R.C., Lopes J.M., Pereira R., Tallon B.E.,
 RA Hatfield G.W.;
 RT "The complete nucleotide sequence of the *ilvGMDA* operon of
 RT *Escherichia coli* K-12.";
 RL Nucleic Acids Res. 15:2137-2155(1987).
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=92358234; PubMed=1379743;
 RA Daniels D.U., Plunkett G. III, Burland V.D., Blattner F.R.;
 RT "Analysis of the *Escherichia coli* genome: DNA sequence of the region
 RT from 84.5 to 86.5 minutes.";
 RL Science 257:771-778(1992).
 RN [4]
 RC SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;

DR	EMBL:	X02413;	CAA26262.1;	-
DR	EMBL:	M32253;	AAA24022.1;	-
DR	EMBL:	V00290;	CAA23559.1;	-
DR	PIR:	A00602;	XNCEV.	
DR	PIR:	S30668;	S30668.	
DR	PDB:	1A3G;	27-MAY-98.	
DR	SWISS-2DPAGE:	P00510;	COLI.	
DR	ECCODBASE:	F032.5;	6TH EDITION.	
DR	Ecocore:	EG10497;	llve.	
DR	InterPro:	IPR001344;	Aminotran_4.	
DR	Pfam:	PF01063;	aminotran_4; 1.	
DR	ProDom:	PD001961;	Aminotran_4; 1.	
DR	TIGRFAMs:	TIGR01122;	llve.f. 1.	
DR	PROSITE:	PS00770;	AA.TRANSFER_CLASS_4; 1.	
KW	transferase; Aminotransferase;	Branched-chain amino acid biosynthesis;		
KM	pyridoxal phosphate;	3d-structure; Complete proteome.		
FT	INIT_MET	0		
FT	MOD_RES	1	THR MODIFIED TO GLY.	
FT	BINDING	159	PYRIDOXAL PHOSPHATE.	
FT	CONFLICT	150	A -> R (IN REF. 3).	
SQ	SEQUENCE	308 AA;	33962 MW; 28404978D2B7E959 CRC64;	
Oy	15 QDXKHYYGSXKYGXA 31			
Db	18 EDKRVHMSHALHYGTS 34			
RESULT 15				
ID	llve_salty			
AC	llve_salty	STANDARD;	PRT;	308 AA.
DT	01-APR-1990 (Rel. 14.	Created)		
DT	01-APR-1990 (Rel. 14.	Last sequence update)		
DT	15-JUN-2002 (Rel. 41.	Last annotation update)		
DE	Branched-chain amino acid aminotransferase (EC 2.6.1.42) (Transaminase [BCAT]).			
GN	llve OR STM903 OR SMDL.89 OR STY3654.			
OS	Salmonella typhimurium, and			
OC	Salmonella typh.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Salmonella.			
OX	NCBI_Taxid=602, 601;			
RN	[1]			
RP	SEQUENCE.			
RC	SPECIES-S.typhimurium;			
RX	MEDLINE=89352621; PubMed=2669973;			
RA	Feild M.J., Nguyen D.C., Armstrong F.B.;			
RT	"Amino acid sequence of Salmonella typhimurium branched-chain amino acid aminotransferase.";			
RL	Biochemistry 28:5306-5310(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES-S.typhimurium; STRAIN-LT2 / SGSC1412 / ATCC 700720;			
RA	MCDLind=21534948; PubMed=11677609;			
RA	McClintand M., Sanderson K.E., Spiehl J., Clifton S.W., Latreille P.,			
RA	Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,			
RA	Leonard S., Nguyen C., Scott K., Holmes A., Grevell N., Mulvaney E.,			
RA	Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,			
RA	Waterston R., Wilson R.K.;			
RT	Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.;			
RL	Nature 413:852-856(2001).			
RN	[3]			
RP	SEQUENCE OF 1-5.			
RC	SPECIES-S.typhimurium;			
RA	Randall I.R., Wallis M.H., Young G.J., Armstrong F.B.;			
RT	"N-terminal sequence of branched-chain amino acid aminotransferase.";			
RL	Fed. Proc. 38:325-326(1979).			
RN	[4]			

```

RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Goira P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
CC -1- FUNCTION: Acts on leucine, isoleucine and valine.
CC -1- CATALYTIC ACTIVITY: L-leucine + 2-oxoglutarate -> 4-methyl-2-
CC oxopentanoate + L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Valine and isoleucine biosynthesis.
CC -1- SUBUNIT: HOMOHexamER.
CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF233324; AAF34481.1; -.
DR EMBL: AF008882; AAL22753.1; -.
DR EMBL: AF627279; CAD0944.1; -.
DR PIR: A05077; A05077.
DR PIR: A34082; A34082.
DR HSSP: P00510; 1A3G.
DR StyGene: SG10181; 1IVE.
DR InterPro: IPR001544; Aminotran_4.
DR Pfam: PF01063; aminotran_4; 2.
DR ProDom: PD001961; Aminotran_4; 1.
DR TIGRFAMs: TIGR01122; 1IVE_1; 1.
DR PROSITE: PS00770; AA_TRANSFER_CLASS_4; 1.
KW transferase; Aminotransferase; Branched-chain amino acid biosynthesis;
KW Pyridoxal phosphate; Complete proteome.
FT INT_MET 0
FT BINDING 159 159 PYRIDOXAL PHOSPHATE.
SO SEQUENCE 308 AA; 33921 MW; 8F826434F93688AD CRC64;

```

Query Match 25.7%; Score 45; DB 1; Length 308;
 Best Local Similarity 35.3%; Pred. No. 17;
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 15 ODXKIHYGYSMXYGXA 31
 :|:|:|:::|:|:
 Db 18 EDKVVHVMShALHYGTS 34

Search completed: May 7, 2003, 19:04:01
 Job time : 6.93578 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:04:05 ; Search time 8.02294 Seconds
(without alignments)
378,521 Million cell updates/sec

Title: US-09-914-831-4

Perfect score: 175
Sequence: 1 DCEGLGGRIHQSDPKIHVYGYSMXGXAQH 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 9205710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Published Applications, AA.*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	30.3	494	10	US-09-764-864-1235
2	50	28.6	374	10	US-09-815-242-10571
3	49.5	28.3	120	9	US-09-764-868-772
4	48.5	27.7	397	9	US-09-847-208-17
5	48	27.4	391	10	US-09-815-242-13175
6	48	27.4	656	9	US-09-801-7364-2
7	48	27.4	678	9	US-09-892-598-179
8	48	27.4	678	9	US-09-889-293A-179
9	48	27.4	678	9	US-10-063-547-34
10	48	27.4	678	9	US-09-989-735-179
11	48	27.4	678	9	US-09-990-444-179
12	48	27.4	678	9	US-09-989-730-179
13	48	27.4	678	9	US-09-890-436-179
14	48	27.4	678	9	US-09-891-181-179
15	48	27.4	678	9	US-09-993-687-179
16	48	27.4	678	9	US-09-989-734-179
17	48	27.4	678	9	US-09-997-653-179
18	48	27.4	678	9	US-10-174-590-150
19	48	27.4	678	9	US-10-176-758-150

20	48	27.4	578	9	US-10-063-616-34	Sequence 34, App1
21	48	27.4	578	9	US-10-175-737-150	Sequence 150, App
22	48	27.4	578	9	US-09-993-667-179	Sequence 179, App
23	48	27.4	578	9	US-10-063-502-34	Sequence 34, App1
24	48	27.4	578	9	US-10-173-706-150	Sequence 150, App
25	48	27.4	578	9	US-10-175-738-150	Sequence 150, App
26	48	27.4	578	9	US-10-175-752-150	Sequence 150, App
27	48	27.4	578	9	US-10-176-482-150	Sequence 150, App
28	48	27.4	578	9	US-10-176-757-150	Sequence 150, App
29	48	27.4	578	9	US-10-176-913-150	Sequence 150, App
30	48	27.4	578	9	US-10-180-552-150	Sequence 150, App
31	48	27.4	578	9	US-10-180-557-150	Sequence 150, App
32	48	27.4	578	9	US-09-990-438-179	Sequence 179, App
33	48	27.4	578	9	US-09-990-562-179	Sequence 179, App
34	48	27.4	578	9	US-09-997-666-179	Sequence 179, App
35	48	27.4	578	9	US-10-173-700-150	Sequence 150, App
36	48	27.4	578	9	US-10-174-572-150	Sequence 150, App
37	48	27.4	578	9	US-10-174-579-150	Sequence 150, App
38	48	27.4	578	9	US-10-174-582-150	Sequence 150, App
39	48	27.4	578	9	US-10-174-588-150	Sequence 150, App
40	48	27.4	578	9	US-10-175-739-150	Sequence 150, App
41	48	27.4	578	9	US-10-175-740-150	Sequence 150, App
42	48	27.4	578	9	US-10-175-743-150	Sequence 150, App
43	48	27.4	578	9	US-10-176-488-150	Sequence 150, App
44	48	27.4	578	9	US-10-176-492-150	Sequence 150, App
45	48	27.4	578	9	US-10-176-492-150	Sequence 150, App

ALIGNMENTS

RESULT 1
US-09-764-864-1235
Sequence 1235 Application US/09764864
Patent No. US20020132735A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT23
CURRENT APPLICATION NUMBER: US/09/764,864
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 1235
LENGTH: 494
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (14)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (488)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (493)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (494)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1235
Query Match
Best Local Similarity 30.3% Score 53; DB 10; Length 494;
Matches 9; Conservative 4; Mismatches 10; Gaps 0;
Oy 4 CUGGRIHQSDPKIHVYGYSM 26
Db 407 CTPNAEVREHQDGTTHLYYSI 429
RESULT 2

US-09-815-242-10571
; Sequence 10571, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10571
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10571

Query Match 28.6%; Score 50; DB 10; Length 374;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 ECLGGGRISHQSODXK 18
||||| | | | |
Db 354 ECLGGGLIDRAYQDEK 369

RESULT 3
US-09-764-868-772
; Sequence 772, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 772
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-772

Query Match 28.3%; Score 49.5; DB 9; Length 120;
Best Local Similarity 35.5%; Pred. No. 4.5;
Matches 11; Conservative 6; Mismatches 11; Indels 3; Gaps 1;

QY 2 CECIGGRISHOSODKIHVYSGMXGXAQ 32
| | | | | : : : | | |
Db 81 CQVSSREAIHSFQELQVHILG---NWGLAQ 108

RESULT 4
US-09-847-208-17
; Sequence 17, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: UC67,002A
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Ambrosia artemisiifolia (Short ragweed)
US-09-847-208-17

Query Match 27.7%; Score 48.5; DB 9; Length 397;
Best Local Similarity 41.2%; Pred. No. 24;
Matches 14; Conservative 0; Mismatches 11; Indels 9; Gaps 1;

QY 1 DCECLGGGRIS-----HOSODKXIHVYGS 25
| | | | | | | | | | | | | | |
Db 171 DVRVLPGRIRKSNCGPAIRHQSDGDAIHVTGSS 204

RESULT 5
US-09-815-242-13175
; Sequence 13175, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13175
LENGTH: 391
TYPE: PRF
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13175

Query Match 27.4%; Score 48; DB 10; Length 391;
Best Local Similarity 45.0%; Pred. No. 28;
Matches 9; Conservativity 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 ECLGGRIHQSDXKIHVY 22
||||| : : : :
Db 371 ECLGGGLIDNAYRDGVCQY 390

RESULT 6
US-09-801-736A-2
Sequence 2, Application US/09801736A
Patent No. US20020160937A1
GENERAL INFORMATION:
APPLICANT: Mayne, Richard
Ren, Zhao-Xia
Liu, Jiangang
TITLE OF INVENTION: Human Vitreous and Pharmaceutical Compositions
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glenna Hendricks
STREET: P.O. Box 2509
CITY: Fairfax
STATE: VA
COUNTRY: U.S.A.
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/801,736A
FILING DATE: 09-Mar-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: uab-may
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 425-8405
TELEFAX: (703) 425-8406
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 656 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
TISSUE TYPE: amino acid sequence of human vitreous humor
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-801-736A-2

Query Match 27.4%; Score 48; DB 9; Length 656;
Best Local Similarity 48.0%; Pred. No. 50;
Matches 12; Conservativity 0; Mismatches 7; Indels 6; Gaps 1;

QY 15 QDPKHVGTDTYTSYSSVCAAVH 91
||||| : : : :
Db 67 QDPKHVGTDTYTSYSSVCAAVH 91

RESULT 7
US-09-992-598-179
Sequence 179, Application US/09992598
Patent No. US20020160384A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Iver J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, F. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same
FILE REFERENCE: P2730P1C20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948

PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.4%; Score 48; DB 9; Length 678;
Best Local Similarity 48.0%; Pred. No. 51;
Matches 12; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

OY 15 QDXKHIVG-----YSKXGYGAOH 33

DB 67 QDPKHYHGTDTYASTVSCGAHVH 91

RESULT 8

US-09-989-293A-179

Sequence 179, Application US/09989293A

Patent No. US2002017716A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William T.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC66
CURRENT APPLICATION NUMBER: US/09/989,293A
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22

PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.4%; Score 48; DB 9; Length 678;
Best Local Similarity 48.0%; Pred. No. 51;
Matches 12; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

QY 15 ODKKHVYG-----YSMXGXAOH 33
DB 67 QDPKHYVGTDTYASTSYSCGAAYH 91

RESULT 9
US-10-063-547-34
; Sequence 34, Application US/10063547

Publication No. US20020182638A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,547
CURRENT FILING DATE: 2002-05-02
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 34
LENGTH: 678
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-547-34

Query Match 27.4%; Score 48; DB 9; Length 678;
Best Local Similarity 48.0%; Pred. No. 51;
Matches 12; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

QY 15 ODKKHVYG-----YSMXGXAOH 33
DB 67 QDPKHYVGTDTYASTSYSCGAAYH 91

RESULT 10
US-09-989-735-179
; Sequence 179, Application US/09989735
; Publication No. US20020193299A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deasoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C61
CURRENT APPLICATION NUMBER: US/09/989,735
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12

1	PRIOR APPLICATION NUMBER: 60/065311
2	PRIOR FILING DATE: 1997-11-13
3	PRIOR APPLICATION NUMBER: 60/066770
4	PRIOR FILING DATE: 1997-11-24
5	PRIOR APPLICATION NUMBER: 60/075945
6	PRIOR FILING DATE: 1998-02-25
7	PRIOR APPLICATION NUMBER: 60/078910
8	PRIOR FILING DATE: 1998-03-20
9	PRIOR APPLICATION NUMBER: 60/083322
10	PRIOR FILING DATE: 1998-04-28
11	PRIOR APPLICATION NUMBER: 60/084600
12	PRIOR FILING DATE: 1998-05-07
13	PRIOR APPLICATION NUMBER: 60/087106
14	PRIOR FILING DATE: 1998-05-28
15	PRIOR APPLICATION NUMBER: 60/087607
16	PRIOR FILING DATE: 1998-06-02
17	PRIOR APPLICATION NUMBER: 60/087609
18	PRIOR FILING DATE: 1998-06-02
19	PRIOR APPLICATION NUMBER: 60/087759
20	PRIOR FILING DATE: 1998-06-02
21	PRIOR APPLICATION NUMBER: 60/087827
22	PRIOR FILING DATE: 1998-06-03
23	PRIOR APPLICATION NUMBER: 60/088021
24	PRIOR FILING DATE: 1998-06-04
25	PRIOR APPLICATION NUMBER: 60/088025
26	PRIOR FILING DATE: 1998-06-04
27	PRIOR APPLICATION NUMBER: 60/088026
28	PRIOR FILING DATE: 1998-06-04
29	PRIOR APPLICATION NUMBER: 60/088028
30	PRIOR FILING DATE: 1998-06-04
31	PRIOR APPLICATION NUMBER: 60/088029
32	PRIOR FILING DATE: 1998-06-04
33	PRIOR APPLICATION NUMBER: 60/088030
34	PRIOR FILING DATE: 1998-06-04
35	PRIOR APPLICATION NUMBER: 60/088033
36	PRIOR FILING DATE: 1998-06-04
37	PRIOR APPLICATION NUMBER: 60/088326
38	PRIOR FILING DATE: 1998-06-04
39	PRIOR APPLICATION NUMBER: 60/088167
40	PRIOR FILING DATE: 1998-06-05
41	PRIOR APPLICATION NUMBER: 60/088202
42	PRIOR FILING DATE: 1998-06-05
43	PRIOR APPLICATION NUMBER: 60/088212
44	PRIOR FILING DATE: 1998-06-05
45	PRIOR APPLICATION NUMBER: 60/088217
46	PRIOR FILING DATE: 1998-06-05
47	PRIOR APPLICATION NUMBER: 60/088655
48	PRIOR FILING DATE: 1998-06-09
49	PRIOR APPLICATION NUMBER: 60/088734
50	PRIOR FILING DATE: 1998-06-10
51	PRIOR APPLICATION NUMBER: 60/088738
52	PRIOR FILING DATE: 1998-06-10
53	PRIOR APPLICATION NUMBER: 60/088742
54	PRIOR FILING DATE: 1998-06-10
55	PRIOR APPLICATION NUMBER: 60/088810
56	PRIOR FILING DATE: 1998-06-10
57	PRIOR APPLICATION NUMBER: 60/088824
58	PRIOR FILING DATE: 1998-06-10
59	PRIOR APPLICATION NUMBER: 60/088826
60	PRIOR FILING DATE: 1998-06-10
61	PRIOR APPLICATION NUMBER: 60/088858
62	PRIOR FILING DATE: 1998-06-11
63	PRIOR APPLICATION NUMBER: 60/088861
64	PRIOR FILING DATE: 1998-06-11
65	PRIOR APPLICATION NUMBER: 60/088876
66	PRIOR FILING DATE: 1998-06-11
67	PRIOR APPLICATION NUMBER: 60/089105
68	PRIOR FILING DATE: 1998-06-12
69	PRIOR APPLICATION NUMBER: 60/089440
70	PRIOR FILING DATE: 1998-06-16
71	PRIOR APPLICATION NUMBER: 60/089512
72	PRIOR FILING DATE: 1998-06-16
73	PRIOR APPLICATION NUMBER: 60/089514

[illegible]

PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/0833322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/0844600
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/0871066
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/0876077
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/0876099
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/0877959
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/0878277
PRIOR FILING DATE:	1998-06-03
PRIOR APPLICATION NUMBER:	60/0880211
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880255
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880266
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880282
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880299
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880300
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0880333
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0883262
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/0884166
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/0882022
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/0882121
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/0882171
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/0885555
PRIOR FILING DATE:	1998-06-09
PRIOR APPLICATION NUMBER:	60/0887344
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0887383
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0887424
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888100
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888222
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888222
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/0888555
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/0888666
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/0888777
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/0891055
PRIOR FILING DATE:	1998-06-12
PRIOR APPLICATION NUMBER:	60/0894444
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/0895121
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/0895146
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/0895146
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/0895333
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/0895999
PRIOR FILING DATE:	1998-06-17

PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090442
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090555
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/09057
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626

PRIOR FILING DATE: 1993-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1993-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1993-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1993-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1993-07-09

Query Match 27.4%; Score 48; DB 9; Length 678;
Best Local Similarity 48.0%; Pred. No. 51;
Matches 12; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

QY 15 QDXKHVYG-----YSMXGXAOH 33
Db 67 QDPKHVHYGIDVYASYSWCGAAVH 91

RESULT 12
US-09-989-730-179
; Sequence 179, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC69
; CURRENT APPLICATION NUMBER: US/09/989,730
; PRIOR APPLICATION NUMBER: 2001-11-20
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28

1	PRIOR APPLICATION NUMBER: 60/087607
2	PRIOR FILING DATE: 1998-06-02
3	PRIOR APPLICATION NUMBER: 60/087609
4	PRIOR FILING DATE: 1998-06-02
5	PRIOR APPLICATION NUMBER: 60/087759
6	PRIOR FILING DATE: 1998-06-02
7	PRIOR APPLICATION NUMBER: 60/088021
8	PRIOR FILING DATE: 1998-06-04
9	PRIOR APPLICATION NUMBER: 60/088025
10	PRIOR FILING DATE: 1998-06-04
11	PRIOR APPLICATION NUMBER: 60/088026
12	PRIOR FILING DATE: 1998-06-04
13	PRIOR APPLICATION NUMBER: 60/088028
14	PRIOR FILING DATE: 1998-06-04
15	PRIOR APPLICATION NUMBER: 60/088029
16	PRIOR FILING DATE: 1998-06-04
17	PRIOR APPLICATION NUMBER: 60/088030
18	PRIOR FILING DATE: 1998-06-04
19	PRIOR APPLICATION NUMBER: 60/088033
20	PRIOR FILING DATE: 1998-06-04
21	PRIOR APPLICATION NUMBER: 60/088326
22	PRIOR FILING DATE: 1998-06-04
23	PRIOR APPLICATION NUMBER: 60/088167
24	PRIOR FILING DATE: 1998-06-05
25	PRIOR APPLICATION NUMBER: 60/088202
26	PRIOR FILING DATE: 1998-06-05
27	PRIOR APPLICATION NUMBER: 60/088212
28	PRIOR FILING DATE: 1998-06-05
29	PRIOR APPLICATION NUMBER: 60/088217
30	PRIOR FILING DATE: 1998-06-05
31	PRIOR APPLICATION NUMBER: 60/088655
32	PRIOR FILING DATE: 1998-06-09
33	PRIOR APPLICATION NUMBER: 60/088734
34	PRIOR FILING DATE: 1998-06-10
35	PRIOR APPLICATION NUMBER: 60/088738
36	PRIOR FILING DATE: 1998-06-10
37	PRIOR APPLICATION NUMBER: 60/088742
38	PRIOR FILING DATE: 1998-06-10
39	PRIOR APPLICATION NUMBER: 60/088810
40	PRIOR FILING DATE: 1998-06-10
41	PRIOR APPLICATION NUMBER: 60/088824
42	PRIOR FILING DATE: 1998-06-10
43	PRIOR APPLICATION NUMBER: 60/088826
44	PRIOR FILING DATE: 1998-06-10
45	PRIOR APPLICATION NUMBER: 60/088858
46	PRIOR FILING DATE: 1998-06-11
47	PRIOR APPLICATION NUMBER: 60/088861
48	PRIOR FILING DATE: 1998-06-11
49	PRIOR APPLICATION NUMBER: 60/088876
50	PRIOR FILING DATE: 1998-06-11
51	PRIOR APPLICATION NUMBER: 60/089105
52	PRIOR FILING DATE: 1998-06-12
53	PRIOR APPLICATION NUMBER: 60/089440
54	PRIOR FILING DATE: 1998-06-16
55	PRIOR APPLICATION NUMBER: 60/089512
56	PRIOR FILING DATE: 1998-06-16
57	PRIOR APPLICATION NUMBER: 60/089514
58	PRIOR FILING DATE: 1998-06-16
59	PRIOR APPLICATION NUMBER: 60/089532
60	PRIOR FILING DATE: 1998-06-17
61	PRIOR APPLICATION NUMBER: 60/089538
62	PRIOR FILING DATE: 1998-06-17
63	PRIOR APPLICATION NUMBER: 60/089598
64	PRIOR FILING DATE: 1998-06-17
65	PRIOR APPLICATION NUMBER: 60/089599
66	PRIOR FILING DATE: 1998-06-17
67	PRIOR APPLICATION NUMBER: 60/089600
68	PRIOR FILING DATE: 1998-06-17
69	PRIOR APPLICATION NUMBER: 60/089653
70	PRIOR FILING DATE: 1998-06-17
71	PRIOR APPLICATION NUMBER: 60/089601

1	PRIOR FILING DATE: 1998-06-18
2	PRIOR APPLICATION NUMBER: 60/089907
3	PRIOR FILING DATE: 1998-06-18
4	PRIOR APPLICATION NUMBER: 60/089908
5	PRIOR FILING DATE: 1998-06-18
6	PRIOR APPLICATION NUMBER: 60/089947
7	PRIOR FILING DATE: 1998-06-19
8	PRIOR APPLICATION NUMBER: 60/089948
9	PRIOR FILING DATE: 1998-06-19
10	PRIOR APPLICATION NUMBER: 60/089952
11	PRIOR FILING DATE: 1998-06-19
12	PRIOR APPLICATION NUMBER: 60/090246
13	PRIOR FILING DATE: 1998-06-22
14	PRIOR APPLICATION NUMBER: 60/090252
15	PRIOR FILING DATE: 1998-06-22
16	PRIOR APPLICATION NUMBER: 60/090254
17	PRIOR FILING DATE: 1998-06-22
18	PRIOR APPLICATION NUMBER: 60/090349
19	PRIOR FILING DATE: 1998-06-23
20	PRIOR APPLICATION NUMBER: 60/090355
21	PRIOR FILING DATE: 1998-06-23
22	PRIOR APPLICATION NUMBER: 60/090429
23	PRIOR FILING DATE: 1998-06-24
24	PRIOR APPLICATION NUMBER: 60/090431
25	PRIOR FILING DATE: 1998-06-24
26	PRIOR APPLICATION NUMBER: 60/090435
27	PRIOR FILING DATE: 1998-06-24
28	PRIOR APPLICATION NUMBER: 60/090444
29	PRIOR FILING DATE: 1998-06-24
30	PRIOR APPLICATION NUMBER: 60/090445
31	PRIOR FILING DATE: 1998-06-24
32	PRIOR APPLICATION NUMBER: 60/090472
33	PRIOR FILING DATE: 1998-06-24
34	PRIOR APPLICATION NUMBER: 60/090555
35	PRIOR FILING DATE: 1998-06-24
36	PRIOR APPLICATION NUMBER: 60/090540
37	PRIOR FILING DATE: 1998-06-24
38	PRIOR APPLICATION NUMBER: 60/090542
39	PRIOR FILING DATE: 1998-06-24
40	PRIOR APPLICATION NUMBER: 60/090557
41	PRIOR FILING DATE: 1998-06-24
42	PRIOR APPLICATION NUMBER: 60/090676
43	PRIOR FILING DATE: 1998-06-25
44	PRIOR APPLICATION NUMBER: 60/090678
45	PRIOR FILING DATE: 1998-06-25
46	PRIOR APPLICATION NUMBER: 60/090690
47	PRIOR FILING DATE: 1998-06-25
48	PRIOR APPLICATION NUMBER: 60/090654
49	PRIOR FILING DATE: 1998-06-25
50	PRIOR APPLICATION NUMBER: 60/090655
51	PRIOR FILING DATE: 1998-06-25
52	PRIOR APPLICATION NUMBER: 60/090656
53	PRIOR FILING DATE: 1998-06-25
54	PRIOR APPLICATION NUMBER: 60/090862
55	PRIOR FILING DATE: 1998-06-26
56	PRIOR APPLICATION NUMBER: 60/090863
57	PRIOR FILING DATE: 1998-06-26
58	PRIOR APPLICATION NUMBER: 60/091360
59	PRIOR FILING DATE: 1998-07-01
60	PRIOR APPLICATION NUMBER: 60/091519
61	PRIOR FILING DATE: 1998-07-02
62	PRIOR APPLICATION NUMBER: 60/091626
63	PRIOR FILING DATE: 1998-07-02
64	PRIOR APPLICATION NUMBER: 60/091633
65	PRIOR FILING DATE: 1998-07-02
66	PRIOR APPLICATION NUMBER: 60/091978
67	PRIOR FILING DATE: 1998-07-07
68	PRIOR APPLICATION NUMBER: 60/091982
69	PRIOR FILING DATE: 1998-07-07

PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089807
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252

PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.4%; Score 48; DB 9; Length 678;
Best Local Similarity 48.0%; Pred. No. 51;
Matches 12; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

QY 15 ODXKHVYG-----YSKXGXAQH 33
Db 67 QDPKHYHGTDVYASVSCGAHVH 91

RESULT 15
US-09-993-687-179

Sequence 179, Application US/09993687
Publication No. US20020198149A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C11
CURRENT APPLICATION NUMBER: US/09/993,687
CURRENT FILING DATE: 2002-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033

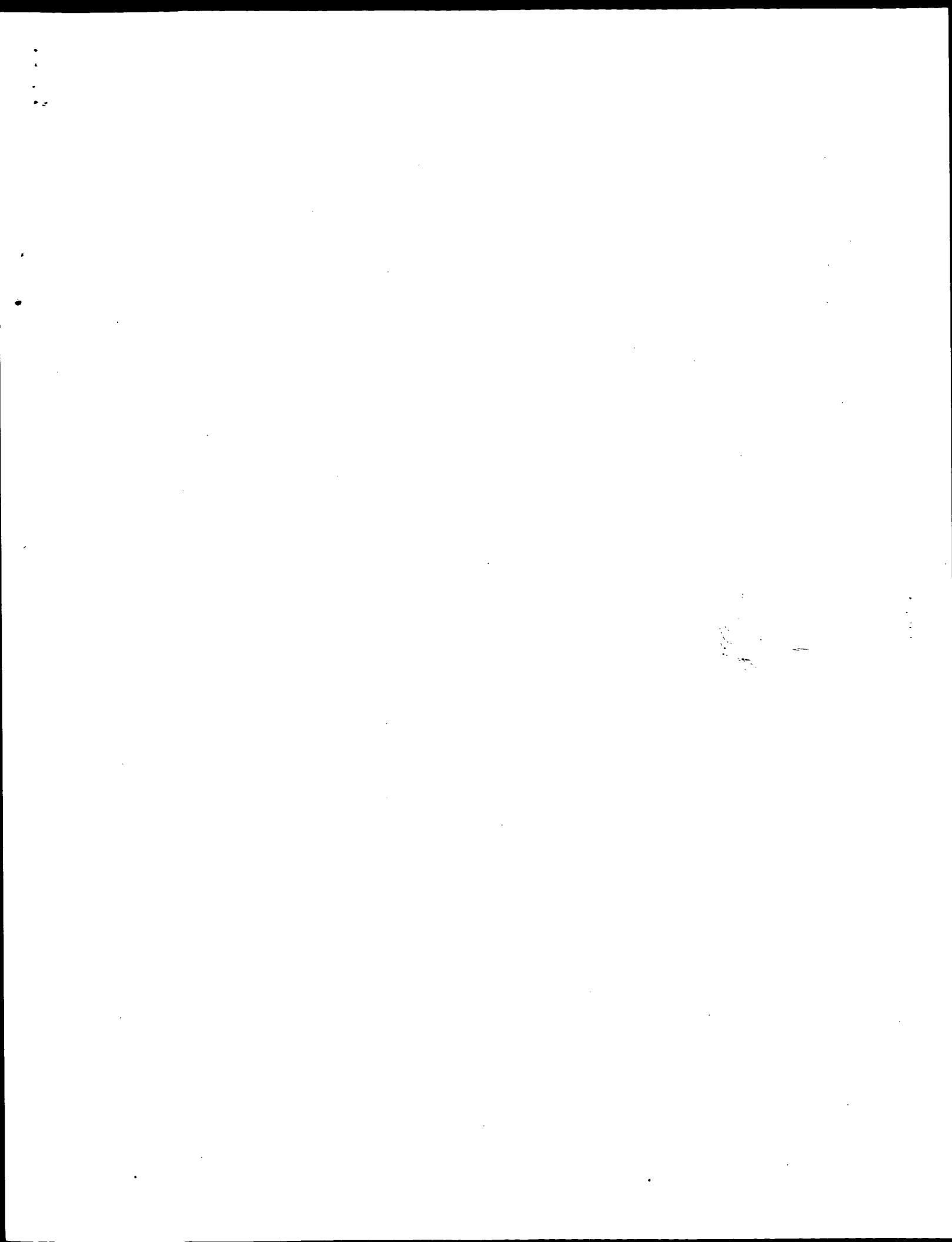
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23

1	PRIOR APPLICATION NUMBER: 60/090424
2	PRIOR FILING DATE: 1998-06-24
3	PRIOR APPLICATION NUMBER: 60/090431
4	PRIOR FILING DATE: 1998-06-24
5	PRIOR APPLICATION NUMBER: 60/090435
6	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090444
8	PRIOR FILING DATE: 1998-06-24
9	PRIOR APPLICATION NUMBER: 60/090445
10	PRIOR FILING DATE: 1998-06-24
11	PRIOR APPLICATION NUMBER: 60/090472
12	PRIOR FILING DATE: 1998-06-24
13	PRIOR APPLICATION NUMBER: 60/090535
14	PRIOR FILING DATE: 1998-06-24
15	PRIOR APPLICATION NUMBER: 60/090540
16	PRIOR FILING DATE: 1998-06-24
17	PRIOR APPLICATION NUMBER: 60/090542
18	PRIOR FILING DATE: 1998-06-24
19	PRIOR APPLICATION NUMBER: 60/090557
20	PRIOR FILING DATE: 1998-06-24
21	PRIOR APPLICATION NUMBER: 60/090676
22	PRIOR FILING DATE: 1998-06-25
23	PRIOR APPLICATION NUMBER: 60/090678
24	PRIOR FILING DATE: 1998-06-25
25	PRIOR APPLICATION NUMBER: 60/090690
26	PRIOR FILING DATE: 1998-06-25
27	PRIOR APPLICATION NUMBER: 60/090694
28	PRIOR FILING DATE: 1998-06-25
29	PRIOR APPLICATION NUMBER: 60/090695
30	PRIOR FILING DATE: 1998-06-25
31	PRIOR APPLICATION NUMBER: 60/090696
32	PRIOR FILING DATE: 1998-06-25
33	PRIOR APPLICATION NUMBER: 60/090862
34	PRIOR FILING DATE: 1998-06-26
35	PRIOR APPLICATION NUMBER: 60/090863
36	PRIOR FILING DATE: 1998-06-26
37	PRIOR APPLICATION NUMBER: 60/091360
38	PRIOR FILING DATE: 1998-07-01
39	PRIOR APPLICATION NUMBER: 60/091478
40	PRIOR FILING DATE: 1998-07-02
41	PRIOR APPLICATION NUMBER: 60/091544
42	PRIOR FILING DATE: 1998-07-01
43	PRIOR APPLICATION NUMBER: 60/091519
44	PRIOR FILING DATE: 1998-07-02
45	PRIOR APPLICATION NUMBER: 60/091626
46	PRIOR FILING DATE: 1998-07-02
47	PRIOR APPLICATION NUMBER: 60/091633
48	PRIOR FILING DATE: 1998-07-02
49	PRIOR APPLICATION NUMBER: 60/091978
50	PRIOR FILING DATE: 1998-07-07
51	PRIOR APPLICATION NUMBER: 60/091982
52	PRIOR FILING DATE: 1998-07-07
53	PRIOR APPLICATION NUMBER: 60/092182
54	PRIOR FILING DATE: 1998-07-09

Query Match 27.4%; Score 48; DB 9; Length 678;
Best Local Similarity 48.0%; Pred. NO. 51;
Matches 12; Conservative 0; Mismatches 7; Indels 6; Caps 1;

```
QY 15 QDXKHVYG-----YSMXYGXAQH 33
      || | |||      || | |
Db 67 QDPKYHVYGTVDVYASYSVCGAANH 91
```

Search completed: May 7, 2003, 19:18:49
Job time : 10.0229 secs



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:04:05 ; Search time 3.88991 Seconds

(without alignments)
378,521 Million cell updates/sec

Title: US-09-914-831-3

Perfect score: 92

Sequence: 1 DCECLGGGRISHOSQD 16

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	50.0	374	10	US-09-815-242-10571
2	44	47.8	43	10	US-09-864-761-38634
3	44	47.8	353	12	US-10-071-751-65
4	44	47.8	353	12	US-10-071-751-68
5	44	47.8	599	9	US-10-113-852A-2
6	44	47.8	641	9	US-09-900-449A-8
7	44	47.8	774	9	US-09-900-449A-2
8	43	46.7	391	10	US-09-815-242-11375
9	42	45.7	164	10	US-09-925-301-1270
10	42	45.7	231	10	US-09-925-302-358
11	42	45.7	239	9	US-09-940-235-4
12	42	45.7	586	9	US-10-092-390-4
13	42	45.7	1140	9	US-10-092-390-2
14	42	45.7	2386	9	US-09-961-403-1
15	41	44.6	298	10	US-09-815-242-11172
16	40	43.5	149	9	US-09-991-496-86
17	40	43.5	149	10	US-09-874-923-86
18	39	42.4	155	10	US-09-864-761-45830
19	39	42.4	1174	10	US-09-826-660-6

20	39	42.4	1177	9	US-10-035-060-2	Sequence 2, Appl
21	39	42.4	1177	9	US-10-035-060-6	Sequence 6, Appl
22	39	42.4	1177	9	US-10-035-060-8	Sequence 8, Appl
23	39	42.4	1177	10	US-09-873-873-10	Sequence 10, Appl
24	39	42.4	1177	10	US-09-873-873-12	Sequence 12, Appl
25	39	42.4	1177	10	US-09-873-873-14	Sequence 14, Appl
26	39	42.4	1177	10	US-09-873-873-16	Sequence 16, Appl
27	39	42.4	1177	10	US-09-873-873-18	Sequence 18, Appl
28	39	42.4	1177	10	US-09-873-873-20	Sequence 20, Appl
29	39	42.4	1177	10	US-09-873-873-22	Sequence 22, Appl
30	39	42.4	1177	10	US-09-873-873-24	Sequence 24, Appl
31	39	42.4	1177	10	US-09-873-873-26	Sequence 26, Appl
32	39	42.4	1177	10	US-09-873-873-28	Sequence 28, Appl
33	39	42.4	1177	10	US-09-873-873-30	Sequence 30, Appl
34	39	42.4	1177	10	US-09-873-873-32	Sequence 32, Appl
35	39	42.4	1177	10	US-09-873-873-34	Sequence 34, Appl
36	39	42.4	1177	10	US-09-873-873-36	Sequence 36, Appl
37	39	42.4	1177	10	US-09-873-873-38	Sequence 38, Appl
38	39	42.4	1177	10	US-09-873-873-40	Sequence 40, Appl
39	39	42.4	1177	10	US-09-873-873-42	Sequence 42, Appl
40	39	42.4	1177	10	US-09-873-873-44	Sequence 44, Appl
41	39	42.4	1177	10	US-09-873-873-46	Sequence 46, Appl
42	39	42.4	1177	10	US-09-873-873-48	Sequence 48, Appl
43	39	42.4	1177	10	US-09-873-873-50	Sequence 50, Appl
44	39	42.4	1177	10	US-09-873-873-52	Sequence 52, Appl
45	39	42.4	1177	10	US-09-873-873-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-10571
Sequence 10571, Application US/09815242
Patient No. US2002061569A1

GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10571
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10571

Query Match 50.0%; Score 46; DB 10; Length 374;
Best Local Similarity 64.3%; Pred. No. 22;

Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 ECLGGGRSHOSQ 16
|||||
Db 354 ECLGGGLDRAYD 367

RESULT 2
US-09-864-761-38634
; Sequence 38634, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38634
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005477.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: P11956, EVALUATE 4.80e+00

US-09-864-761-38634

Query Match 47.8%; Score 44; DB 10; Length 43;
Best Local Similarity 60.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 DCECLGGGRSHOSQ 15
|||||
Db 14 DCECLCMDGASHOCQ 28

RESULT 3
US-10-071-751-65
; Sequence 65, Application US/10071751
; Patent No. US20020142352A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Slim, Gek-kee
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
; APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHERIDAN ROSS P.C.
; STREET: 1560 BROADWAY, SUITE 1200
; CITY: DENVER
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/071,751
; FILING DATE: 07-Feb-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/171,156
; FILING DATE: 1998-10-09
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-071-751-65
Query Match 47.8%; Score 44; DB 12; Length 353;
Best Local Similarity 46.7%; Pred. No. 42;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 ECLGGGRSHOSQ 16
|||:|:|:|:|
Db 17 CDYGGGPKTHKSD 31

RESULT 4
US-10-071-751-68
; Sequence 68, Application US/10071751
; Patent No. US20020142352A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Slim, Gek-kee

```

?      Weber, Eric R.
?      TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
?      APPARATUS TO COLLECT SUCH PROTEINS
?      NUMBER OF SEQUENCES: 88
?      CORRESPONDENCE ADDRESS:
?      ADDRESSEE: SHERIDAN ROSS P.C.
?      STREET: 1560 BROADWAY, SUITE 1200
?      CITY: DENVER
?      STATE: CO
?      COUNTRY: U.S.A.
?      ZIP: 80202
?      COMPUTER READABLE FORM:
?      MEDIUM TYPE: Floppy disk
?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: PatentIn Release #1.0, Version #1.30
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/10/071,751
?      FILING DATE: 07-Feb-2002
?      CLASSIFICATION: <Unknown>
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: 09/171,156
?      FILING DATE: 1998-10-09
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Connell, Gary J.
?      REGISTRATION NUMBER: 32,020
?      REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 303/863-9700
?      TELEFAX: 303/863-0223
?      INFORMATION FOR SEQ ID NO: 68:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 353 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      SEQUENCE DESCRIPTION: SEQ ID NO: 68:
?      US-10-071-751-68

Query Match          47.8%; Score 44; DB 12; Length 353;
Best Local Similarity 72.7%; Pred. No. 42;
Matches              7; Conservative           5; Mismatches       3; Indels     0; Gaps    0;

Cy      2 CECGGGRIHQSD 16
        1: 11:11:11:1
Db      17 CDYPGGPKTRHSED 31

RESULT 5
US-10-113-852A-2
: Sequence 2, Application US/10113852A
: Publication No. US20020192784A1
: GENERAL INFORMATION:
: APPLICANT: ApplIng, Dean R.
: APPLICANT: Hanson, Andrew D.
: APPLICANT: Raymond, Rhonda R.
: APPLICANT: Roje, Sanja
: TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine in a Recombinant Yeast Strain
: FILE REFERENCE: 119927-1033
: CURRENT APPLICATION NUMBER: US/10/113,852A
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: 60/280,333
: PRIOR FILING DATE: 2001-03-30
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 599
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
: US-10-113-852A-2

```

```

Matches      8; Conservative      1; Mismatches      2; Indels      0; Gaps      0;
QY      6 GCGRISHQSD 16
      |||||:||||
Db      58 GGGRI$HSLSTD 68

RESULT 6
US-09-900-449A-8
; Sequence 8, Application US/09900449A
; Publication No. US20030040616A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Jenny et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01271
; CURRENT APPLICATION NUMBER: US/09/900,449A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 641
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-449A-8

Query Match
Best Local Similarity 47.8%; Score 44; DB 9; Length 641;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 DCECLGGGR 9
      |||||
Db      105 DCECRNGGR 113

RESULT 7
US-09-900-449A-2
; Sequence 2, Application US/09900449A
; Publication No. US20030040616A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Jenny et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01271
; CURRENT APPLICATION NUMBER: US/09/900,449A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-449A-2

Query Match
Best Local Similarity 47.8%; Score 44; DB 9; Length 774;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 DCECLGGGR 9
      |||||
Db      206 DCECRNGGR 214

RESULT 8
US-09-815-242-13175
; Sequence 13175, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.

```

```

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13175
LENGTH: 391
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13175

```

```

Query Match
Best Local Similarity 46.7%; Score 43; DB 10; Length 391;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 3 ECLGGGRISHQSD 16
||||| 1 : 1
Db 371 ECLGGGLIDNAYRD 384

```

```

RESULT 9
US-09-925-301-1270
Sequence 1270, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1270
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (13)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (138)

```

```

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (152)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (161)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (164)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1270

```

```

Query Match
Best Local Similarity 45.7%; Score 42; DB 10; Length 164;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 DCECLGGR 9
||| 1 : 1
Db 41 DCTCIGAGR 49

```

```

RESULT 10
US-09-925-302-548
Sequence 548, Application US/09925302
Patent No. US2002004941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 548
LENGTH: 231
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (44)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (205)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (212)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (226)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-548

```

```

Query Match
Best Local Similarity 45.7%; Score 42; DB 10; Length 231;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 DCECLGGR 9
||| 1 : 1
Db 151 DCTCIGAGR 159

```

```

RESULT 11
US-09-940-235-4
Sequence 4, Application US/09940235
Publication No. US20030059921A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait

```


APPLICANT: Rajagopal, Kammar
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 259
TYPE: PRT
ORGANISM: Homo sapiens
US-09-940-235-4

Query Match 45.7%; Score 42; DB 9; Length 259;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCEICGGR 9
||:|:|
DB 91 DCTCIGGR 99

RESULT 12
US-10-092-390-4
Sequence 4, Application US/10092390
Publication No. US20030013865A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Miranda, Maricar
TITLE OF INVENTION: No. US20030013865A1 Human EGF-Family Proteins and Polynucleotid
FILE REFERENCE: LEX-0317-USA
CURRENT APPLICATION NUMBER: US/10/092,390
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: US 60/275,013
PRIOR FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 586
TYPE: PRT
ORGANISM: homo sapiens
US-10-092-390-4

Query Match 45.7%; Score 42; DB 9; Length 586;
Best Local Similarity 46.2%; Pred. No. 1,4e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 CECIGGRISHOS 14
||:|:|:|
DB 324 CQCVNGSKCYHVS 336

RESULT 13
US-10-092-390-2
Sequence 2, Application US/10092390
Publication No. US20030013865A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Miranda, Maricar
TITLE OF INVENTION: No. US20030013865A1 Human EGF-Family Proteins and Polynucleotid
FILE REFERENCE: LEX-0317-USA
CURRENT APPLICATION NUMBER: US/10/092,390
CURRENT FILING DATE: 2002-03-06

PRIOR APPLICATION NUMBER: US 60/275,013
PRIOR FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1140
TYPE: PRT
ORGANISM: homo sapiens
US-10-092-390-2

Query Match 45.7%; Score 42; DB 9; Length 1140;
Best Local Similarity 46.2%; Pred. No. 2,6e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 CECIGGRISHOS 14
||:|:|:|
DB 324 CQCVNGSKCYHVS 336

RESULT 14
US-09-961-403-1
Sequence 1, Application US/09961403
Publication No. US20030077589A1
GENERAL INFORMATION:
APPLICANT: HE-STUMP, HOLGER
APPLICANT: HAENDLER, BERNARD
APPLICANT: KRAETZSCHMAR, JOERN
APPLICANT: KREFT, BERTHOLD
APPLICANT: WINTERHAGER, ELKE
APPLICANT: REGIDOR, PEDRO
TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
FILE REFERENCE: SCH-1789
CURRENT APPLICATION NUMBER: US/09/961,403
CURRENT FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-961-403-1

Query Match 45.7%; Score 42; DB 9; Length 2386;
Best Local Similarity 66.7%; Pred. No. 5,4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCEICGGR 9
||:|:|
DB 122 DCTCIGGR 130

RESULT 15
US-09-815-242-11172
Sequence 11172, Application US/09815242
Patent No. US2002061565A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11172
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11172

Query Match 44.68; Score 41; DB 10; Length 298;
Best Local Similarity 53.38; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 CECIGGRISHOSD 16
|||:||||:|
DB 21 CEFQGEWQISHQEKD 35

Search completed: May 7, 2003, 19:18:47
Job time : 4.88991 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:03:30 ; Search time 40.0734 Seconds

(Without alignments)
257.421 Million cell updates/sec

Title: US-09-914-831-3

Sequence: 1 DCECLGGGRISHOSD 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	92	100.0	16	US-09-914-831-3	Sequence 3, Appli
2	92	100.0	33	US-09-914-831-4	Sequence 4, Appli
3	92	100.0	44	US-09-914-831-5	Sequence 5, Appli
4	92	100.0	114	US-09-621-976-6804	Sequence 6804, Ap
5	92	100.0	114	US-60-147-499-6804	Sequence 6804, Ap
6	92	100.0	123	US-09-914-831-7	Sequence 7, Appli

7	92	100.0	124	US-09-914-831-6	Sequence 6, Appli
8	92	100.0	124	US-09-914-831-8	Sequence 8, Appli
9	92	100.0	125	US-09-468-725A-2806	Sequence 2806, Ap
10	92	100.0	125	US-09-621-976-6812	Sequence 6812, Ap
11	92	100.0	125	US-09-914-831-2	Sequence 2, Appli
12	92	100.0	125	US-60-147-499-6812	Sequence 6812, Ap
13	92	100.0	143	US-09-834-366-20093	Sequence 20093, A
14	92	100.0	143	US-60-197-873-20093	Sequence 20093, A
15	92	100.0	165	US-09-488-725A-6378	Sequence 6378, A
16	92	100.0	165	US-09-622-746A-109	Sequence 109, App
17	92	100.0	165	US-09-758-461-619	Sequence 619, App
18	92	100.0	271	PCT-US01-08656-8127	Sequence 8127, A
19	92	100.0	271	PCT-US01-14827-11672	Sequence 11672, A
20	92	100.0	271	PCT-US01-14827-12789	Sequence 12789, A
21	92	95.7	144	US-09-758-461-569	Sequence 569, App
22	73	79.3	98	US-09-621-976-6811	Sequence 6811, App
23	73	79.3	98	US-60-147-499-6811	Sequence 6811, App
24	72	78.3	14	US-09-914-831-10	Sequence 10, Appl
25	48	52.2	144	US-09-675-784A-8285	Sequence 8285, Ap
26	47	51.1	135	US-09-614-150-3084	Sequence 3084, Ap
27	47	51.1	135	US-60-167-217-3138	Sequence 3138, Ap
28	47	51.1	135	US-60-173-464-2576	Sequence 2576, Ap
29	47	51.1	135	US-60-191-637-3090	Sequence 3090, Ap
30	47	51.1	135	US-60-191-681-2478	Sequence 2478, Ap
31	47	51.1	709	US-60-161-932-1231	Sequence 1231, Ap
32	46	50.0	48	US-08-070-153-22	Sequence 22, Appl
33	46	50.0	48	US-08-280-288-22	Sequence 22, Appl
34	46	50.0	48	US-08-280-288A-22	Sequence 22, Appl
35	46	50.0	366	US-09-791-537-120717	Sequence 120717, A
36	46	50.0	374	PCT-US02-03987-10571	Sequence 10571, A
37	46	50.0	374	US-09-815-242-10571	Sequence 10571, A
38	46	50.0	374	US-10-072-851-10571	Sequence 10571, A
39	45	48.9	130	PCT-US01-01324-1543	Sequence 1543, Ap
40	45	48.9	130	US-10-079-979-1543	Sequence 1543, Ap
41	45	48.9	402	US-09-791-537-13166	Sequence 13166, A
42	44	47.8	43	PCT-US01-00663-11825	Sequence 31825, A
43	44	47.8	43	US-09-864-761-38634	Sequence 38634, A
44	44	47.8	43	US-10-182-993-30872	Sequence 30872, A
45	44	47.8	25	US-10-182-995-25106	Sequence 25106, A

ALIGNMENTS

RESULT 1
US-09-914-831-3
Sequence 3, Application US/09914831
GENERAL INFORMATION:
APPLICANT: Merck Patent GmbH
TITLE OF INVENTION: Histidine Protein Phosphatase
FILE REFERENCE: Histidine Phosphatase Protein
CURRENT APPLICATION NUMBER: US/09/914,831
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 16
TYPE: PRT
ORGANISM: mammalian
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(16)
OTHER INFORMATION: conserved mammalian sequence

Query Match 100.0%; Score 92; DB 23; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DCECLGGGRISHOSD 16
Db 1 DCECLGGGRISHOSD 16

RESULT 2
US-09-914-831-4
; Sequence 4, Application US/09914831
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Histidine Protein Phosphatase
; FILE REFERENCE: Histidine Phosphatase Protein
; CURRENT APPLICATION NUMBER: US/09/914,831
; CURRENT FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 33
; TYPE: PRT
; ORGANISM: mammalian
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)-(33)
; OTHER INFORMATION: conserved mammalian sequence 2
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: X = K or R
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: X = A or G
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: X = P or R
US-09-914-831-4

Query Match
Best Local Similarity 100.0%; Score 92; DB 23; Length 33;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
Db 1 DCECLGGGRISHOSOD 16
RESULT 3
US-09-914-831-5
; Sequence 5, Application US/09914831
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Histidine Protein Phosphatase
; FILE REFERENCE: Histidine Phosphatase Protein
; CURRENT APPLICATION NUMBER: US/09/914,831
; CURRENT FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 44
; TYPE: PRT
; ORGANISM: mammalian
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)-(44)
; OTHER INFORMATION: conserved mammalian sequence 3
US-09-914-831-5

Query Match
Best Local Similarity 100.0%; Score 92; DB 23; Length 44;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
Db 19 DCECLGGGRISHOSOD 34
RESULT 4
US-09-621-976-6804
; Sequence 6804, Application US/09621976

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6804
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 106
; OTHER INFORMATION: Xaa = Ser,Thr
US-09-621-976-6804

Query Match
Best Local Similarity 100.0%; Score 92; DB 20; Length 114;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
Db 70 DCECLGGGRISHOSOD 85
RESULT 5
US-60-147-499-6804
; Sequence 6804, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6804
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 106
; OTHER INFORMATION: Xaa = Ser,Thr
US-60-147-499-6804

Query Match
Best Local Similarity 100.0%; Score 92; DB 27; Length 114;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
Db 70 DCECLGGGRISHOSOD 85
RESULT 6
US-09-914-831-7
; Sequence 7, Application US/09914831
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Histidine Protein Phosphatase
; FILE REFERENCE: Histidine Phosphatase Protein
; CURRENT APPLICATION NUMBER: US/09/914,831
; CURRENT FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7

LENGTH: 123
TYPE: PRT
ORGANISM: rat
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)-(123)
OTHER INFORMATION: rat histidine protein phosphatase
US-09-914-831-7

Query Match
Best Local Similarity 100.0%; Score 92; DB 23; Length 123;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16
|||||
DB 68 DCECLGGGRISHOSOD 83

RESULT 7
US-09-914-831-6
Sequence 6, Application US/09914831
GENERAL INFORMATION:
APPLICANT: Merck Patent GmbH
TITLE OF INVENTION: Histidine Protein Phosphatase
FILE REFERENCE: Histidine Phosphatase Protein
CURRENT APPLICATION NUMBER: US/09/914,831
CURRENT FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 124
TYPE: PRT
ORGANISM: rabbit
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)-(124)
OTHER INFORMATION: rabbit histidine protein phosphatase
US-09-914-831-6

Query Match
Best Local Similarity 100.0%; Score 92; DB 23; Length 124;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16
|||||
DB 69 DCECLGGGRISHOSOD 84

RESULT 8
US-09-914-831-8
Sequence 8, Application US/09914831
GENERAL INFORMATION:
APPLICANT: Merck Patent GmbH
TITLE OF INVENTION: Histidine Protein Phosphatase
FILE REFERENCE: Histidine Phosphatase Protein
CURRENT APPLICATION NUMBER: US/09/914,831
CURRENT FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 124
TYPE: PRT
ORGANISM: mouse
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)-(124)
OTHER INFORMATION: mouse histidine protein phosphatase
US-09-914-831-8

Query Match
Best Local Similarity 100.0%; Score 92; DB 23; Length 124;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16
|||||
DB 69 DCECLGGGRISHOSOD 84

RESULT 9
US-09-488-725A-2806
Sequence 2806, Application US/09488725A
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/09/488,725A
CURRENT FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: US/09/488,725
PRIORITY FILING DATE: 2000-01-21
PRIORITY APPLICATION NUMBER: US09/552,317
PRIORITY FILING DATE: 2000-04-25
PRIORITY APPLICATION NUMBER: US09/598,042
PRIORITY FILING DATE: 2000-06-20
PRIORITY APPLICATION NUMBER: US09/620,312
PRIORITY FILING DATE: 2000-07-19
PRIORITY APPLICATION NUMBER: US09/653,450
PRIORITY FILING DATE: 2000-08-31
PRIORITY APPLICATION NUMBER: US09/662,191
PRIORITY FILING DATE: 2000-09-14
PRIORITY APPLICATION NUMBER: US09/693,036
PRIORITY FILING DATE: 2000-10-19
PRIORITY APPLICATION NUMBER: US09/727,344
PRIORITY FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7144
SOFTWARE: PL_FL_genes_versions 1.0
SEQ ID NO 2806
LENGTH: 126
TYPE: PRT
ORGANISM: Homo sapiens
US-09-488-725A-2806

Query Match
Best Local Similarity 100.0%; Score 92; DB 18; Length 125;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16
|||||
DB 70 DCECLGGGRISHOSOD 85

RESULT 10
US-09-621-976-6812
Sequence 6812, Application US/09621976
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 6812
LENGTH: 125
TYPE: PRT
ORGANISM: Homo sapiens
US-09-621-976-6812

Query Match
Best Local Similarity 100.0%; Score 92; DB 20; Length 125;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16
|||||
DB 70 DCECLGGGRISHOSOD 85

RESULT 11
US-09-914-831-2
; Sequence 2, Application US/09914831
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Histidine Protein Phosphatase
; FILE REFERENCE: Histidine Phosphatase Protein
; CURRENT APPLICATION NUMBER: US/09/914,831
; CURRENT FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-831-2

Query Match 100.0%; Score 92; DB 23; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
|||||

Db 70 DCECLGGGRISHOSOD 85

RESULT 12
US-60-147-499-6812
; Sequence 6812, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6812
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-147-499-6812

Query Match 100.0%; Score 92; DB 27; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
|||||

Db 70 DCECLGGGRISHOSOD 85

RESULT 13
US-09-834-366-20093
; Sequence 20093, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejain, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm

; SEQ ID NO 20093
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-20093

Query Match 100.0%; Score 92; DB 22; Length 143;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
|||||

Db 70 DCECLGGGRISHOSOD 85

RESULT 14
US-60-197-873-20093
; Sequence 20093, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejain, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 20093
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-197-873-20093

Query Match 100.0%; Score 92; DB 27; Length 143;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHOSOD 16
|||||

Db 70 DCECLGGGRISHOSOD 85

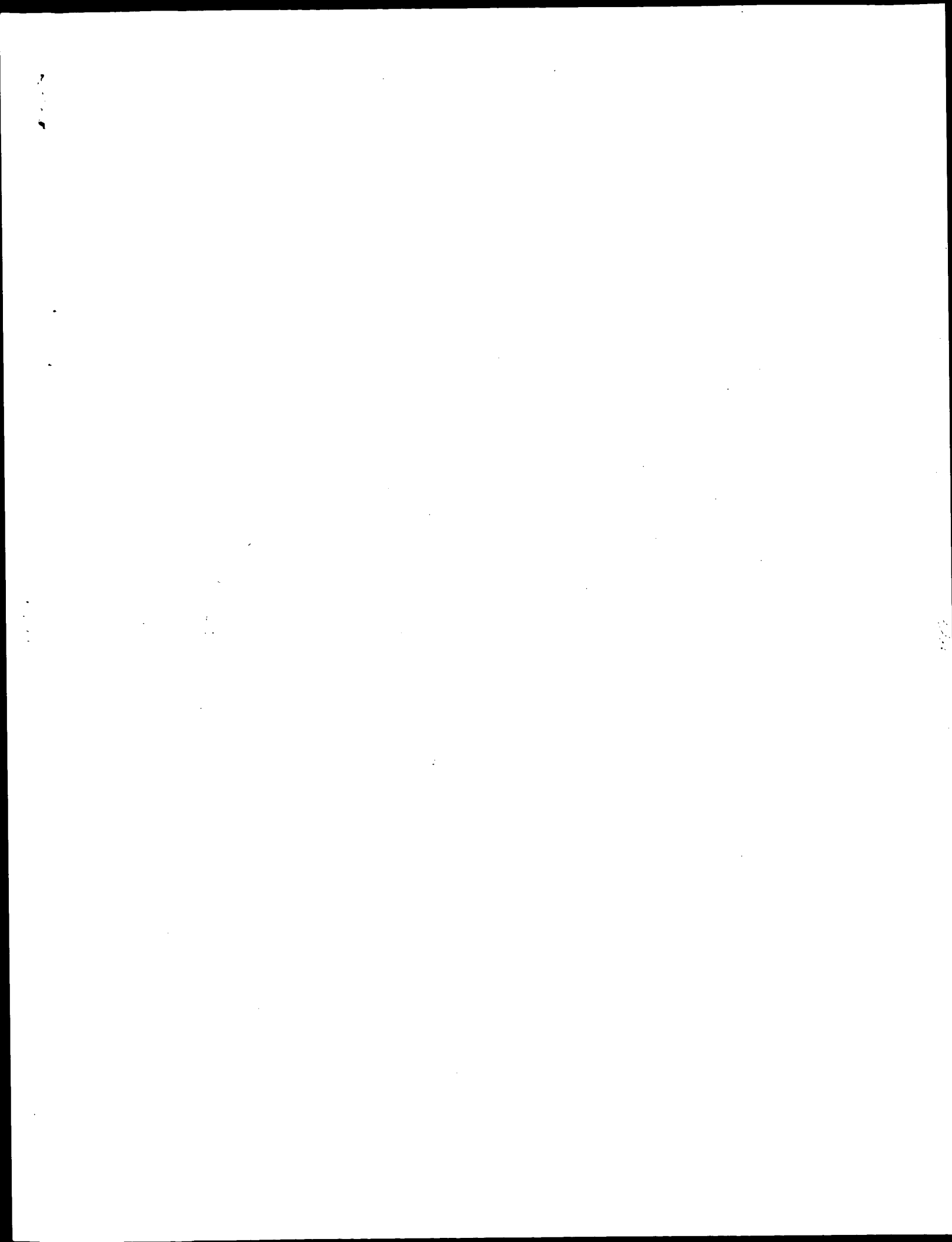
RESULT 15
US-09-488-725A-6378
; Sequence 6378, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hysq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 6378

; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-6378

Query Match 100.0%; Score 92; DB 18; Length 165;
Best Local Similarity 100.0%; Pred. No. 56-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCECLGGGGRISHOSQD 16
|||||
Db 110 DCECLGGGGRISHOSQD 125

Search completed: May 7, 2003, 19:17:47
Job time: 41.0734 secs



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 18:54:20 ; Search time 8.51376 Seconds

(without alignments)
250.419 Million cell updates/sec

Title: US-09-914-831-3
Sequence: 1 DCECIGGRSHSQD 16

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	92	100.0	16	21	AA08788	Conserved sequence
2	92	100.0	33	21	AA08789	Conserved sequence
3	92	100.0	44	21	AA08790	Conserved sequence
4	92	100.0	123	21	AA08792	Conserved sequence
5	92	100.0	124	21	AA08791	Rat histidine prot
6	92	100.0	124	21	AA08793	Rabbit histidine p
7	92	100.0	125	21	AA08787	Mouse histidine pr
8	92	100.0	125	22	AA039661	A human histidine
9	92	100.0	165	20	AA48548	Human polypeptide
10	92	100.0	165	21	AA42776	Human prostate can
						Human ORFX ORF25401

11	92	100.0	165	22	AA041447	Human polypeptide
12	72	78.3	14	21	AA08795	Antigenic peptide
13	47	51.1	135	21	AB058764	Drosophila melanog
14	46	50.0	48	22	AA097795	N-terminal fragmen
15	46	50.0	374	22	AA034978	Enterococcus faeca
16	45	48.9	130	22	AA092194	Human digestive sy
17	44	47.8	43	22	AB038130	Peptide #536 enco
18	44	47.8	43	22	AB023336	Protein #535 enco
19	44	47.8	43	22	AA058767	Human brain expres
20	44	47.8	43	22	AA071275	Human bone marrow
21	44	47.8	43	22	AA031556	Peptide #593 enco
22	44	47.8	43	23	AB041075	Human peptide enco
23	44	47.8	100	23	AB009930	Human ORFX protein
24	44	47.8	353	18	AA030484	Flea saliva protei
25	44	47.8	353	19	AA082374	Flea saliva protei
26	44	47.8	548	22	AA064737	Brugia malayi aspa
27	44	47.8	701	22	AB066725	Drosophila melanog
28	44	47.8	1028	22	AB082249	Human insulin-resp
29	43	46.7	119	22	AA031440	Novel human secret
30	43	46.7	152	19	AA086166	S. pneumoniae deri
31	43	46.7	373	21	AA094213	S. pneumoniae pneu
32	43	46.7	373	21	AA055634	S. pneumoniae pneu
33	43	46.7	373	23	AB030558	Streptococcus poly
34	43	46.7	391	22	AB037582	Streptococcus pneu
35	43	46.7	397	23	AB027008	Streptococcus pneu
36	42	45.7	102	21	AA034510	Human protein SE0
37	42	45.7	164	21	AA033825	Human cancer assoc
38	42	45.7	231	21	AA058210	Lung cancer associ
39	42	45.7	255	21	AA044449	Arabidopsis thalia
40	42	45.7	259	21	AA090281	Human fibronectin
41	42	45.7	259	21	AA044451	Arabidopsis thalia
42	42	45.7	269	21	AA044450	Novel human diagno
43	42	45.7	321	22	AB027639	Herbicide activ
44	42	45.7	366	23	AB091046	Herbicide activ
45	42	45.7	366	23	AB091047	Herbicide activ

ALIGNMENTS

RESULT 1
ID AAB08788 standard. Peptide: 16 AA.
XX AAB08788;
AC 02-JAN-2001 (first entry)
XX
DE Conserved sequence of mammalian histidine protein phosphatases.
XX
KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Mammalia sp.
XX
XX
XX WO200052175-A1.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-EP01774.
XX
XX 04-MAR-1999; 99DE-1009388.
XX
XX (MERE) MERCK PATENT GMBH.
XX
XX Klump S, Kellner R;
XX
XX WPI, 2000-572187/53.
XX
XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
XX
XX immune disorders, viral infection, genetic disorders, and heart disease
XX
XX

XX Claim 2; Page 17; 39pp; English.
XX
XX The present sequence represents a conserved sequence found in mammalian
CC histidine phosphatase polypeptides. The human polypeptide has a high
CC specificity for phosphohistidine and a molecular weight of 13000-15000
CC Da. The human histidine phosphatase gene is localised at chromosome 9
CC (9q33). The human histidine phosphatase can be used for diagnosis and
CC treatment of pathological states of cell regulation and cell growth.
CC These include cancers, immune disorders, viral infection, genetic
CC disorders, and heart disease. The histidine phosphatase can also be
CC used for identifying agonists and antagonists which can be used to
CC treat conditions associated with N-phosphorylation imbalance.
XX
XX Sequence 16 AA:
SQ
Query Match 100.0%; Score 92; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DCECLGGGRISHOSOD 16
DB 1 DCECLGGGRISHOSOD 16
|||||
RESULT 2
ID AAB08789 standard; Peptide; 33 AA.
XX AAB08789;
XX
XX 02-JAN-2001 (first entry)
XX
XX Conserved sequence of mammalian histidine protein phosphatases.
DE
XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
XX Mammalia sp.
OS
XX Key Location/Qualifiers
FH MISC-difference 17
FT /label= Lys, Arg
FT MISC-difference 27
FT /label= Ala, Gly
FT MISC-difference 30
FT /label= Pro, Arg
XX
XX WO200052175-A1.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-EP01774.
XX
XX 04-MAR-1999; 99DE-1009388.
XX
XX (MERE) MERCK PATENT GMBH.
XX
XX Klump S, Kellner R;
XX
XX WPI; 2000-572187/53.
XX
XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
PT
XX
XX Claim 3; Page 17; 39pp; English.
XX
XX The present sequence represents a conserved sequence found in mammalian
CC histidine phosphatase polypeptides. The human polypeptide has a high
CC specificity for phosphohistidine and a molecular weight of 13000-15000
CC Da. The human histidine phosphatase gene is localised at chromosome 9
CC (9q33). The human histidine phosphatase can be used for diagnosis and
CC treatment of pathological states of cell regulation and cell growth.
CC These include cancers, immune disorders, viral infection, genetic
CC disorders, and heart disease. The histidine phosphatase can also be
CC used for identifying agonists and antagonists which can be used to
CC treat conditions associated with N-phosphorylation imbalance.
XX
XX Sequence 16 AA:
SQ
Query Match 100.0%; Score 92; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DCECLGGGRISHOSOD 16
DB 1 DCECLGGGRISHOSOD 16
|||||

CC (9q33). The human histidine phosphatase can be used for diagnosis and
CC treatment of pathological states of cell regulation and cell growth.
CC These include cancers, immune disorders, viral infection, genetic
CC disorders, and heart disease. The histidine phosphatase can also be
CC used for identifying agonists and antagonists which can be used to
CC treat conditions associated with N-phosphorylation imbalance.
XX
XX Sequence 33 AA:
SQ
Query Match 100.0%; Score 92; DB 21; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DCECLGGGRISHOSOD 16
DB 1 DCECLGGGRISHOSOD 16
|||||
RESULT 3
ID AAB08790 standard; Peptide; 44 AA.
XX AAB08790;
XX
XX 02-JAN-2001 (first entry)
XX
XX Conserved sequence of mammalian histidine protein phosphatases.
DE
XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
XX Mammalia sp.
OS
XX WO200052175-A1.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-EP01774.
XX
XX 04-MAR-1999; 99DE-1009388.
XX
XX (MERE) MERCK PATENT GMBH.
XX
XX Klump S, Kellner R;
XX
XX WPI; 2000-572187/53.
XX
XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
PT
XX
XX Claim 4; Page 17; 39pp; English.
XX
XX The present sequence represents a conserved sequence found in mammalian
CC histidine phosphatase polypeptides. The human polypeptide has a high
CC specificity for phosphohistidine and a molecular weight of 13000-15000
CC Da. The human histidine phosphatase gene is localised at chromosome 9
CC (9q33). The human histidine phosphatase can be used for diagnosis and
CC treatment of pathological states of cell regulation and cell growth.
CC These include cancers, immune disorders, viral infection, genetic
CC disorders, and heart disease. The histidine phosphatase can also be
CC used for identifying agonists and antagonists which can be used to
CC treat conditions associated with N-phosphorylation imbalance.
XX
XX Sequence 44 AA:
SQ
Query Match 100.0%; Score 92; DB 21; Length 44;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DCECLGGGRISHOSOD 16
DB 1 DCECLGGGRISHOSOD 16
|||||

Db 19 DCECLGGGRISHOSOD 34

RESULT 4
AAB08792
ID AAB08792 standard; Peptide; 123 AA.
XX
AC AAB08792;
XX
DT 02-JAN-2001 (first entry)
XX
DE Rat histidine protein phosphatase polypeptide.
XX
KW human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Rattus sp.
XX
PN WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
PR 04-MAR-1999; 99DE-1009388.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;
XX
DR WPI; 2000-572187/53.
XX
PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
XX
PS Disclosure; Page 12; 39pp; English.
XX
CC The present sequence represents a rat histidine phosphatase
CC polypeptide. The human polypeptide has a high specificity for
CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
CC histidine phosphatase gene is localised at chromosome 9 (9q33). The
CC human histidine phosphatase can be used for diagnosis and treatment
CC of pathological states of cell regulation and cell growth. These
CC include cancers, immune disorders, viral infection, genetic disorders,
CC and heart disease. The histidine phosphatase can also be used for
CC identifying agonists and antagonists which can be used to treat
CC conditions associated with N-phosphorylation imbalance.
XX
SQ Sequence 123 AA;
XX
Query Match 100.0%; Score 92; DB 21; Length 123;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 DCECLGGGRISHOSOD 16
Db 68 DCECLGGGRISHOSOD 83

RESULT 5
AAB08791
ID AAB08791 standard; Peptide; 124 AA.
XX
AC AAB08791;
XX
DT 02-JAN-2001 (first entry)
XX
DE Rabbit histidine protein phosphatase polypeptide.
XX
KW human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;

KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Oryctolagus cuniculus.
XX
PN WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
PR 04-MAR-1999; 99DE-1009388.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;
XX
DR WPI; 2000-572187/53.
XX
PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
XX
PS Claim 7; Page 18; 39pp; English.
XX
CC The present sequence represents a rabbit histidine phosphatase
CC polypeptide. The human polypeptide has a high specificity for
CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
CC histidine phosphatase gene is localised at chromosome 9 (9q33). The
CC human histidine phosphatase can be used for diagnosis and treatment
CC of pathological states of cell regulation and cell growth. These
CC include cancers, immune disorders, viral infection, genetic disorders,
CC and heart disease. The histidine phosphatase can also be used for
CC identifying agonists and antagonists which can be used to treat
CC conditions associated with N-phosphorylation imbalance.
XX
SQ Sequence 124 AA;
XX
Query Match 100.0%; Score 92; DB 21; Length 124;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 DCECLGGGRISHOSOD 16
Db 69 DCECLGGGRISHOSOD 84

RESULT 6
AAB08793
ID AAB08793 standard; Peptide; 124 AA.
XX
AC AAB08793;
XX
DT 02-JAN-2001 (first entry)
XX
DE Mouse histidine protein phosphatase polypeptide.
XX
KW human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Mus sp.
XX
PN WO200052175-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-EP01774.
XX
PR 04-MAR-1999; 99DE-1009388.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Klump S, Kellner R;

XX DR WPI: 2000-572187/53.
 XX CC Histidine phosphatase, useful for diagnosis and treatment of cancers,
 PT immune disorders, viral infection, genetic disorders, and heart disease
 PT
 XX
 PS Disclosure: Page 12; 39pp: English.
 XX
 CC The present sequence represents a mouse histidine phosphatase
 CC polypeptide. The human polypeptide has a high specificity for
 CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
 CC histidine phosphatase gene is located at chromosome 9 (9q33). The
 CC human histidine phosphatase can be used for diagnosis and treatment
 CC of pathological states of cell regulation and cell growth. These
 CC include cancers, immune disorders, viral infection, genetic disorders,
 CC and heart disease. The histidine phosphatase can also be used for
 CC identifying agonists and antagonists which can be used to treat
 CC conditions associated with N-phosphorylation imbalance.
 XX
 SQ Sequence 124 AA;
 Query Match 100.0%; Score 92; DB 21; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DCECLGGGRISHOSQD 16
 DB 69 DCECLGGGRISHOSQD 84
 RESULT 7
 AAB08787 standard; Protein; 125 AA.
 AAB08787:
 02-JAN-2001 (first entry)
 A human histidine protein phosphatase polypeptide.
 Human: histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 cell regulation; cell growth; cancer; immune disorder; viral infection;
 genetic disorder; heart disease; N-phosphorylation imbalance.
 Homo sapiens.
 WO200052175-A1.
 08-SEP-2000.
 02-MAR-2000; 2000WO-EP01774.
 04-MAR-1999; 99DE-1009388.
 (MERE) MERCK PATENT GMBH.
 Klump S, Kellner R;
 WPI: 2000-572187/53.
 N-PSDB: AAA64502.
 Histidine phosphatase, useful for diagnosis and treatment of cancers,
 immune disorders, viral infection, genetic disorders, and heart disease
 Histidine phosphatase, useful for diagnosis and treatment of cancers,
 immune disorders, viral infection, genetic disorders, and heart disease
 Claim 5; Page 17; 39pp: English.
 The present sequence represents human histidine phosphatase polypeptide.
 The polypeptide has a high specificity for phosphohistidine and a
 molecular weight of 13000-15000 Da. The histidine phosphatase gene is
 localised at chromosome 9 (9q33). The histidine phosphatase can be used
 for diagnosis and treatment of pathological states of cell regulation

CC and cell growth. These include cancers, immune disorders, viral
 CC infection, genetic disorders, and heart disease. The histidine
 CC phosphatase can also be used for identifying agonists and antagonists
 CC which can be used to treat conditions associated with N-phosphorylation
 CC imbalance.
 CC
 SQ Sequence 125 AA;
 Query Match 100.0%; Score 92; DB 21; Length 125;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DCECLGGGRISHOSQD 16
 DB 70 DCECLGGGRISHOSQD 85
 RESULT 8
 AAM39661 standard; Protein; 125 AA.
 AAM39661:
 22-OCT-2001 (first entry)
 Human polypeptide SPQ ID NO 2806.
 Human: nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 Leukemia.
 Homo sapiens.
 WO200153312-A1.
 26-JUL-2001.
 26-DEC-2000; 2000WO-US34263.
 21-JAN-2000; 2000US-0488725.
 25-APR-2000; 2000US-0552317.
 09-JUL-2000; 2000US-0598042.
 19-JUL-2000; 2000US-0620312.
 03-AUG-2000; 2000US-0653450.
 14-SEP-2000; 2000US-0662191.
 19-OCT-2000; 2000US-0693036.
 29-NOV-2000; 2000US-0727344.
 (HYSE-) HYSEQ INC.
 Tang YT, Liu C, Asundi T, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 Zhao Q, Zhou P, Goodrich R, Drmanac RT;
 WPI: 2001-442253/47.
 N-PSDB: AAI58817.
 Novel nucleic acids and polypeptides, useful for treating disorders
 such as central nervous system injuries -
 Example 4; SEQ ID NO 2806; 10078pp; English.
 The invention relates to human nucleic acids (AAI57798-AAI61369) and
 the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 immunosuppressant and cyostatic activity. The polynucleotides are useful
 in gene therapy. A composition containing a polypeptide or polynucleotide
 of the invention may be used to treat diseases of the peripheral nervous
 system, such as peripheral nervous injuries, peripheral neuropathy and
 localised neuropathies and central nervous system diseases, such as
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC antitumour haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 165 AA;

Query Match 100.0%; Score 92; DB 21; Length 165;

Best Local Similarity 100.0%; Pred. No. 1.3e-05; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16
 DB 110 DCECLGGGRISHOSOD 125

RESULT 11

ID AAM41447 standard; Protein; 165 AA.

AC AAM41447;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6378.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX anisotropic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukemia.

OS Homo sapiens.

PN WO20015312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-0534263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

DR N-PSDB; AAI60603.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PS Example 2; SEQ ID NO 6378; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

XX Specification.

XX Sequence 165 AA;

Query Match 100.0%; Score 92; DB 22; Length 165;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSOD 16

DB 110 DCECLGGGRISHOSOD 125

RESULT 12

ID AAB08795 standard; Peptide; 14 AA.

AC AAB08795;

DT 02-JAN-2001 (first entry)

DE Antigenic peptide derived from histidine protein phosphatase.

XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;

XX cell regulation; cell growth; cancer; immune disorder; viral infection;

XX genetic disorder; heart disease; N-phosphorylation imbalance.

OS Synthetic.

PN WO200052175-A1.

PD 08-SEP-2000.

PF 02-MAR-2000; 2000WO-EP01774.

PR 04-MAR-1999; 99DE-1009388.

PR (MERE) MERCK PATENT GMBH.

PI Klump S, Kellner R;

PI WPI: 2000-572187/53.

PT Histidine phosphatase, useful for diagnosis and treatment of cancers,

PT immune disorders, viral infection, genetic disorders, and heart disease

PS Disclosure; Page 13; 39pp; English.

XX The present sequence represents an antigenic peptide derived from a

CC mammalian histidine phosphatase polypeptide. The peptide is used to

CC raise antibodies. The human polypeptide has a high specificity for

Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CLGGGRISHOSQ 16
DB 1 CLGGGRISHOSQ 13

RESULT 13

ABB58764
ID ABB58764 standard; Protein: 135 AA.

AC ABB58764;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 3084.

KW Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-6556860/75.

XX N-PSDB; ABL02867.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

PS Disclosure; SEQ ID NO 3084; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 135 AA;

XX Query Match 51.1%; Score 47; DB 22; Length 135;

XX Best Local Similarity 61.5%; Pred. No. 26;

XX Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSQ 15

DB 82 ECRGGGRIEHNP 94

RESULT 14

AA97795

XX AA97795 standard; Peptide; 48 AA.

AC AA97795;

XX 22-AUG-2001 (first entry)

DE N-terminal fragment of hUPA.

KW Human; hUPA; urokinase-type plasminogen activator; uPA; therapy;
uPA-mediated disorder; non-fucosylated polypeptide; tumour metastasis;
inappropriate angiogenesis; corneal angiogenesis; chronic inflammation;
Kaposi's sarcoma; diabetic retinopathy; rheumatoid arthritis;
N-terminal fragment.

XX Homo sapiens.

OS US6248715-B1.

XX 19-JUN-2001.

XX 10-MAY-1995; 95US-0438745.

XX 26-JUL-1994; 94US-0280288.

XX 01-JUN-1993; 93US-0070153.

XX (CHIR) CHIRON CORP.

XX Rosenberg S, Stratton-Thomas JR;

XX WPI; 2001-380524/40.

XX Treatment of disorders mediated by urokinase-type plasminogen activator
(UPA), e.g. tumour metastasis and corneal angiogenesis, comprises
administering a recombinant UPA fragment -

XX Disclosure; Column 39-40; 21pp; English.

XX This sequence represents the N-terminus of human UPA (hUPA).

XX The invention relates to a method for treating urokinase-type plasminogen
activator (UPA)-mediated disorders comprising administering a composition
comprising a non-fucosylated polypeptide. The method is useful for
treating tumour metastasis, inappropriate angiogenesis (especially
corneal angiogenesis), chronic inflammation, Kaposi's sarcoma, diabetic
retinopathy and rheumatoid arthritis.

XX Sequence 48 AA;

XX Query Match 50.0%; Score 46; DB 22; Length 48;

XX Best Local Similarity 53.8%; Pred. No. 14;

XX Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 DDECLGGGRISHQ 13

DB 10 NCDCLNGGTVCVHQ 22

RESULT 15

AAU34978

XX AAU34978 standard; Protein: 374 AA.

XX AAU34978;

XX 14-FEB-2002 (first entry)

XX Enterococcus faecalis cellular proliferation protein #265.

XX Antisense; prokaryotic cellular proliferation protein;
antibiotic; antibacterial; drug design.

XX Enterococcus faecalis.

OS WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX PA
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.
 DR N-PSDB; AAS52837.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -

XX PS Example 3; Seq ID No 10571; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC http://wipo.int/pub/published_pct_sequences.

XX SQ Sequence 374 AA;

Query Match 50.0%; Score 46; DB 22; Length 374;
 Best Local Similarity 64.3%; Pred. No. 96;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 ECLGGRIHQSD 16

Db 354 ECLGGGLIDRAYQD 367

Search completed: May 7, 2003, 19:03:25
 Job time : 10.5138 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:00 ; Search time 3.44954 Seconds

(without alignments)
445,900 Million cell updates/sec

Title: US-09-914-831-3

Perfect score: 92

Sequence: 1 DCECLGGGRISHQSD 16

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.4	149	2	S53876	sex-regulated prot
2	51.1	135	2	A32317	sex-regulated prot
3	46	115	2	T21833	hypothetical prote
4	45	394	2	C83840	hypothetical prote
5	44	265	2	S34668	nodulation protein
6	44	548	2	A54510	63k antigen - nema
7	44	548	2	A28209	60k filarial anti
8	44	552	2	T10850	YAP protein - Rhi
9	44	599	2	S64136	methylenetetrahyd
10	43	373	2	G95013	hypothetical prote
11	43	391	2	B97887	tRNA (5-methylam
12	43	444	2	S35191	tubulin beta-1 cha
13	43	467	2	T33531	hypothetical prote
14	43	477	2	T33531	alkaline exonuclea
15	43	584	1	QOHE36	hypothetical prote
16	42	192	1	XMEC1F	regulator of flag
17	42	192	2	A8954	flagellar transcri
18	42	192	2	E85802	regulator of flag
19	42	194	2	A40746	flagellar transcri
20	42	201	2	S61277	flagellar transcri
21	42	367	2	F86315	hypothetical prote
22	42	371	2	E83807	(5-methylaminometh
23	42	441	2	E80136	tubulin beta subun
24	42	457	2	E80136	fibronectin - bovi
25	42	457	2	E80136	fibronectin - precu
26	42	457	2	E80136	fibronectin - precu
27	42	457	2	E80136	fibronectin - precu
28	41	298	2	B64180	atac-like transcri
29	41	312	2	G69978	probable ATPase yr

30	41	44.6	700	2	S57194	calpain (EC 3.4.22
31	41	44.6	1020	2	A29355	fibronectin - chic
32	40.5	44.0	309	2	T08150	chitinase (EC 3.2.
33	40	43.5	134	2	AD0323	Nrd1 protein homol
34	40	43.5	194	2	A10202	flagellum biosynth
35	40	43.5	195	2	F96979	molecular chaperon
36	40	43.5	354	2	JT0594	site-specific DNA-
37	40	43.5	369	2	T18663	hypothetical prote
38	40	43.5	378	2	S72599	hypothetical prote
39	40	43.5	395	2	S42680	phosphoserine tran
40	40	43.5	423	2	D87670	peptidase, M16 fam
41	40	43.5	445	2	A54515	tubulin beta chain
42	40	43.5	907	2	AB1885	hypothetical prote
43	40	43.5	1026	2	G87346	hypothetical prote
44	40	43.5	1166	2	S32645	paraspore crystal
45	39.5	42.9	389	2	I49263	potential oncogene

ALIGNMENTS

RESULT 1

S53876 sex-regulated protein janus A - fruit fly (Drosophila pseudoobscura)

C:Species: Drosophila pseudoobscura

C>Date: 14-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Jan-1999

C:Accession: S53876

R:Yanicoskas, C.; Ferrer, P.; Vincent, A.; Lepesant, J.A.

Mol. Gen. Genet. 246, 549-560, 1995

A:Title: Separate cis-regulatory sequences control expression of serendipity beta and

A:Reference number: S53876; MUID:95214615; PMID:7700229

A:Accession: S53876

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-149 <YAN>

A:Cross-references: GB:S77099; NID:9913977; PID:9913978

C:Genetics:

A:Gene: janA

A:Cross-references: FlyBase:FBgn0015151

A:introns: 78/L; 119/3

Query Match

Best Local Similarity 55.4%; Score 51; DB 2; Length 149;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DCECLGGGRISH 12

DB 94 DTGCLGGGRIGH 105

RESULT 2

A32317 sex-regulated protein janA - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 29-Jan-1990 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000

C:Accession: A32317

R:Yanicoskas, C.; Vincent, A.; Lepesant, J.A.

Mol. Cell. Biol. 9, 2526-2535, 1989

A:Title: Transcriptional and posttranscriptional regulation contributes to the sex-re

A:Reference number: A32317; MUID:89343970; PMID:2503707

A:Accession: A32317

A:Status: preliminary

A:Molecule type: DNA; mRNA

A:Residues: 1-135 <YAN>

A:Cross-references: GB:M27033; NID:93522914; PIDN:AAC34203.1; PID:9387597

C:Note: the authors translated the codon Ttg for residue 98 as Met

C:Genetics:

A:Gene: FlyBase:janA

A:Cross-references: FlyBase:FBgn0001280

Query Match

Best Local Similarity 51.1%; Score 47; DB 2; Length 135;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSQ 15
 |||||
 DB 82 ECPGGGRISHNDE 94

RESULT 3

hypothetical protein F36A2.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21833

R:Submitted to the EMBL Data Library, October 1996

A:Reference number: Z19476

A:Accession: T21833

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-115 <NII>

A:Cross-references: EMBL:Z81077; PIDN:CA03070.1; GSPDB:GN00019; CESP:F36A2.8

A:Experimental source: clone F36A2

C:Genetics:

A:Gene: CESP:F36A2.8

A:Map position: 1

A:Introns: 22/3; 52/3

Query Match

Best Local Similarity 50.0%; Score 46; DB 2; Length 115;
 Best Local Similarity 70.0%; Pred. No. 2;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 |||||
 DB 62 KCVGGGRIRK 71

RESULT 4

hypothetical protein BH1523 [Imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: C83840

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: C83840

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-394 <STO>

A:Cross-references: GB:AP001512; GB:BA000004; NID:910174030; PIDN:BA05242.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1523

Query Match

Best Local Similarity 48.8%; Score 45; DB 2; Length 394;
 Best Local Similarity 54.5%; Pred. No. 9; 6;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 13
 |||||
 DB 32 ECPGGGRIRK 42

RESULT 5

modulation protein nol265 - Rhizobium sp.

C:Species: Rhizobium sp.

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995

C:Accession: S34668

R:Rochepeau, P.; Fellay, R.; Broughton, W.

submitted to the EMBL Data Library, July 1993

A:Reference number: S34667

A:Accession: S34668

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-265 <ROC>
 A:Cross-references: EMBL:X74068

Query Match

Best Local Similarity 47.8%; Score 44; DB 2; Length 265;
 Best Local Similarity 46.7%; Pred. No. 9; 6;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 CECLGGGRISHOSQ 16
 |||||
 DB 138 CTCGGGRISHOSQ 152

RESULT 6

63K antigen - nematode (Brugia malayi)

C:Species: Brugia malayi

C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 16-Jul-1999

C:Accession: A54510

R:Perrine, K.G.; Denker, J.A.; Nilsen, T.W.

Mol. Biochem. Parasitol. 30, 97-104, 1988

A:Title: A multi-copy gene encodes a potentially protective antigen in Brugia malayi.

A:Reference number: A54510; MUID:8828341; PMID:2840577

A:Accession: A54510

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-548 <PER>

A:Cross-references: GB:U03971; NID:9156056; PIDN:AA27852.1; PID:9156057

C:Genetics:

A:Introns: 23/2; 79/3; 140/1; 163/3; 214/1; 276/3; 321/3; 379/3; 433/3; 481/3

C:Superfamily: lysine-tRNA ligase

Query Match

Best Local Similarity 47.8%; Score 44; DB 2; Length 548;
 Best Local Similarity 61.5%; Pred. No. 19;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 CECLGGGRISHOS 14
 |||||
 DB 317 CAVLGGGRISHOS 329

RESULT 7

60K filarial antigen - nematode (Brugia malayi)

C:Species: Brugia malayi

C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 16-Jul-1999

C:Accession: A28209

R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Perrine, K.G.; Denker, J.A.; Nanduri, J

Proc. Natl. Acad. Sci. U.S.A. 85, 3604-3607, 1988

A:Title: Cloning and characterization of a potentially protective antigen in lymphatic

A:Reference number: A28209; MUID:88217950; PMID:3368467

A:Accession: A28209

A:Molecule type: mRNA

A:Residues: 1-548 <NII>

A:Cross-references: EMBL:J03266; NID:9156052; PIDN:AA27849.1; PID:9156053

C:Superfamily: lysine-tRNA ligase

Query Match

Best Local Similarity 47.8%; Score 44; DB 2; Length 548;
 Best Local Similarity 61.5%; Pred. No. 19;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 CECLGGGRISHOS 14
 |||||
 DB 317 CAVLGGGRISHOS 329

RESULT 8

Y4P protein - Rhizobium sp. (strain NGR234) plasmid pNGR234a

C:Species: Rhizobium sp.

A:Variety: strain NGR234

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T10850
 R:Freiberg, C.; Fellay, R.; Bairoch, A.; Broughton, W.J.; Rosenthal, A.; Perret, X.
 Nature 387, 394-401, 1997
 A:Title: Molecular basis of symbiosis between *Rhizobium* and legumes.
 A:Reference number: 214734; MUID:97305956; PMID:9153424
 A:Accession: T10850
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-552 <FRED>
 A:Cross-references: EMBL:AE000077; NID:g2182428; PID:g2182435
 C:Genetics:
 A:Gene: yahp
 A:Genome: plasmid pNGR234a

Query Match 47.8%; Score 44; DB 2; Length 552;
 Best Local Similarity 46.7%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 Oy 2 CECGGGRISHQSD 16
 Db 138 CTCGGGRSHKLGED 152

RESULT 9

S64136

methylenetetrahydrofolate reductase (NADPH2) (EC 1.5.1.20) MET13 - yeast (*Saccharomyces cerevisiae*)
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 03-Jun-2002
 C:Accession: S64136; S53294; S78029; S72026
 R:Cerdan, E.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Tizon, B.; Cadahia, J.L.;
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64134

A:Accession: S64136
 A:Molecule type: DNA
 A:Residues: 1-599 <CER>
 A:Cross-references: EMBL:Z72647; NID:g1322686; PIDN:CAA96833.1; PID:g1322687; MIPS:YGL12
 A:Experimental source: strain S288C
 R:Hosaka, K.; Nakawa, J.; Kodaki, T.; Ishizu, H.; Yamashita, S.
 J. Biochem. 116, 1317-1321, 1994
 A:Title: Cloning and sequence of the SCG3 gene which is required for inositol prototroph
 A:Reference number: JX0365; MUID:95221330; PMID:7706223
 A:Accession: S53294

A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-113; 'KRSS' <HOS>
 A:Cross-references: EMBL:D21200
 R:Kitakawa, M.; Grack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wittma
 Eur. J. Biochem. 245, 449-456, 1997
 A:Title: Identification and characterization of the genes for mitochondrial ribosomal pr
 A:Reference number: S78018; MUID:97296414; PMID:9151978
 A:Accession: S78029

A:Molecule type: protein
 A:Residues: 39-40, 'LA', '43-47', 'A', '49' <KIT>
 A:Note: this protein was identified as ribosomal protein Yml45, mitochondrial
 R:Rizyon, B.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Cadahia, J.L.; Cerdan, E.
 Yeast 12, 1047-1051, 1996
 A:Title: Identification of a putative methylenetetrahydrofolate reductase by sequence an
 A:Reference number: S72026; MUID:97051592; PMID:8896269
 A:Accession: S72026

A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-72, 'R', '74-599' <TIZ>
 A:Cross-references: EMBL:X94106; NID:g1628448; PIDN:CAA63833.1; PID:g1628449
 A:Note: this protein was identified as putative methylenetetrahydrofolate reductase

C:Genetics:
 A:Gene: SGD:MET13; MET11; MRPL45
 A:Cross-references: MIPS:YGL125W; SGD:S0003093
 A:Map position: 7L
 A:Genome: nuclear
 C:Function: <METH>
 A:Description: oxidoreductase

A:Note: this function seems to contradict the other function assigned to this protein

C:Function: <RIB>
 A:Pathway: protein biosynthesis
 A:Note: this function seems to contradict the other function assigned to this protein
 C:Keywords: mitochondrion; oxidoreductase; protein biosynthesis; ribosome

Query Match 47.8%; Score 44; DB 2; Length 599;
 Best Local Similarity 72.7%; Pred. No. 21;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Oy 6 GGGGRISHQSD 16
 Db 58 GGGGRSHLSTND 68

RESULT 10

G95013

hypothetical protein SP0118 [imported] - *Streptococcus pneumoniae* (strain TIGR4)

C:Species: *Streptococcus pneumoniae*
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: G95013

R:Rettlein, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
 el, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001.
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
 A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: G95013

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-373 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK74304.1; PID:g14971585; GSPDB:GN00164; TIGR:
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP0118

C:Superfamily: probable membrane protein YDL033c
 Query Match 46.7%; Score 43; DB 2; Length 373;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Oy 3 ECLGGGRISHQSD 16
 Db 353 ECLGGGLIDNAYRD 366

RESULT 11

B97887

tRNA (5-methylaminomethyl-2-thiouridylylate)-methyltransferase (EC 2.1.1.61) [imported]

C:Species: *Streptococcus pneumoniae*
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: B97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burett, S.; Dehoff, B.S.
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
 Y. P. Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.

A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: B97887

A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-391 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK98926.1; PID:g15457661; GSPDB:GN00174
 C:Genetics:
 A:Gene: trny

C:Superfamily: probable membrane protein YDL033c
 C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 46.7%; Score 43; DB 2; Length 391;
 Best Local Similarity 57.1%; Pred. No. 21;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 ECLGGGRISHOSOD 16
 ||||| | : :
 DB 371 ECLGGGLIDNAVYD 384

RESULT 12

S35191
 tubulin beta-1 chain - fungus (Trichoderma viride)
 C:Species: Trichoderma viride
 C:Date: 13-Jan-1995 #sequence_revision 09-May-1997 #text_change 13-Aug-1999
 C:Accession: S35191; S25553
 R:Goldman, G.H.; Temmerman, W.; Jacobs, D.; Contreras, R.; van Montagu, M.; Herrera-Estr
 Mol. Gen. Genet. 240, 73-80, 1993
 A:Title: A nucleotide substitution in one of the beta-tubulin genes of Trichoderma virid
 A:Reference number: S35191; MUID:93341462; PMID:8341264
 A:Accession: S35191
 A:Molecule type: DNA
 A:Residues: 1-444 <GOR>
 A:Cross-references: EMBL:Z15054
 R:Goldman, G.H.; Temmerman, W.W.; Herrera-Estrella, A.A.; Jacobs, D.D.; Contreras, R.
 submitted to the EMBL Data Library, July 1992
 A:Description: A nucleotide substitution in one of the beta-tubulin genes of Trichoderma
 A:Reference number: S25553
 A:Accession: S25553
 A:Molecule type: DNA
 A:Residues: 1-5, 'SV', 6-31, 'SS', 32-47, 50-444 <GOR>
 A:Cross-references: EMBL:Z15054; NID:95204; PIDN:CAA78764.1; PID:95205
 C:Genetics:
 A:Gene: tubl
 A:Introns: 4/3; 12/3; 21/3; 53/3
 C:Superfamily: tubulin

Query Match 46.7%; Score 43; DB 2; Length 444;
 Best Local Similarity 58.3%; Pred. No. 23;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCECIGGRISH 12
 ||||| | : :
 DB 126 NCECIGGFQITR 137

RESULT 13

T33531
 hypothetical protein F58E1.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T33531
 R:Wamsley, P.; Twyman, B.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of C. elegans cosmid F58E1.
 A:Reference number: Z21365
 A:Accession: T33531
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-477 <NMM>
 A:Cross-references: EMBL:AF098995; PIDN:AAC67482.1; GSPDB:GN00020; CESP:F58E1.4
 A:Experimental source: strain Bristol N2; clone F58E1
 C:Genetics:
 A:Gene: CESP:F58E1.4
 A:Map position: 2
 A:Introns: 32/1; 104/1; 269/3; 424/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein T05A6.4

Query Match 46.7%; Score 43; DB 2; Length 477;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 DCECIGGRISHOSOD 16
 ||||| | : :
 DB 170 DCECGRIVGYLARD 185

RESULT 14

OQEB76
 alkaline exonuclease (EC 3.1.11.-) - human cytomegalovirus (strain AD169)
 N:Alternate names: UL98 protein
 C:Species: human cytomegalovirus, human herpesvirus 5
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 11-Jun-1999
 C:Accession: S09863
 R:Chen, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Hornselli, T
 M.; Barrell, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovi
 A:Reference number: S09749; MUID:90269039; PMID:2161319
 A:Accession: S09863
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-584 <CHE>
 A:Cross-references: EMBL:X17403; NID:959591; PIDN:CAA35334.1; PID:91780877
 A:Note: possible protein-coding frames are given
 A:Note: The DNA sequence was submitted to EMBL, December 1989, in computer-readable f
 C:Superfamily: human cytomegalovirus alkaline exonuclease
 C:Keywords: exonuclease; hydrolyase

Query Match 46.7%; Score 43; DB 1; Length 584;
 Best Local Similarity 63.6%; Pred. No. 30;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 CECLGGGRISH 12
 ||||| | : :
 DB 511 CELLAGRVRH 521

RESULT 15

S35220
 hypothetical protein - Pseudomonas syringae pv. tomato
 C:Species: Pseudomonas syringae pv. tomato
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000
 C:Accession: S35220
 R:Salmeron, J.M.; Staskawicz, B.J.
 Mol. Gen. Genet. 239, 6-16, 1993
 A:Title: Molecular characterization and hrp dependence of the avirulence gene avrPto
 A:Reference number: S35220; MUID:93288007; PMID:8510663
 A:Accession: S35220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-164 <SAL>
 A:Cross-references: GB:L20425; NID:9309852; PIDN:AAA25728.1; PID:9309853

Query Match 45.7%; Score 42; DB 2; Length 164;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 CLGGRISHQ 13
 |||| | : :
 DB 5 CYGGRMAHQ 14

Search completed: May 7, 2003, 19:07:31
 Job time: 5.44954 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 18:54:40 ; Search time 1.90826 Seconds

(Without alignments)
347.763 Million cell updates/sec

Title: US-09-914-831-3

Sequence: 1 DCECLGGGRSHSQSD 16

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	64.1	146	1 JANA_DROPS	P54364 drosophila
2	47	51.1	135	1 JANA_DROME	P20348 drosophila
3	44	47.8	148	1 SYN_BROMA	P10723 brugia mala
4	44	47.8	552	1 Y4HP_RHISN	P50360 rhizobium s
5	44	47.8	599	1 MTHS_YEAST	P53128 saccharomyc
6	43	46.7	446	1 TBB1_TRTVI	P31664 trichoderma
7	43	46.7	584	1 EXON_HOMVA	P16789 human cytom
8	42	45.7	192	1 FLHC_ECOLI	P11165 escherichia
9	42	45.7	192	1 FLHC_SALTY	O52222 salmonella
10	42	45.7	194	1 FLHC_SERMA	O5807 serratia ma
11	42	45.7	194	1 FLHC_XENNE	O948f3 xenorhabdus
12	42	45.7	371	1 TRMU_BACCH	O9Kdf2 bacillus ha
13	42	45.7	1229	1 C1BB_BACTU	O45739 bacillus th
14	42	45.7	1231	1 C1BD_BACT2	O92425 bacillus th
15	42	45.7	2265	1 F1NC_BOVIN	P07589 bos taurus
16	42	45.7	2386	1 F1NC_HUMAN	P02751 homo sapien
17	42	45.7	2477	1 F1NC_RAT	P04937 rattus norv
18	42	45.7	2481	1 F1NC_XENLA	O91740 xenopus lae
19	41.5	45.1	416	1 KUFF_HUMAN	O9uth9 homo sapien
20	41	44.6	298	1 YAS2_HAEIN	P45008 haemophilus
21	41	44.6	370	1 TRMU_BACST	O95020 bacillus su
22	41	44.6	700	1 CAN2_CHICK	O92178 gallus gall
23	41	44.6	1256	1 F1NC_CHICK	P11722 gallus gall
24	40	43.5	134	1 NRDI_YERPE	O824c6 yersinia pe
25	40	43.5	193	1 FLHC_YEREN	O86047 yersinia en
26	40	43.5	354	1 MT02_HERAD	P25265 herpetosiph
27	40	43.5	375	1 TRMU_CAEEL	O17440 caenorhabdi
28	40	43.5	378	1 DNJ2_MYCLE	O49762 mycobacteri
29	40	43.5	395	1 SERC_YEAST	P33330 saccharomyc
30	40	43.5	444	1 TBB2_DAUCA	O39697 daucus caro
31	40	43.5	445	1 TBB1_LETME	P21146 leishmania
32	40	43.5	1166	1 C1GA_BACTU	O45746 bacillus th
33	39.5	42.9	389	1 WNIB_HUMAN	O00744 homo sapien

34	39.5	42.9	389	1 WNIB_MOUSE	P48614 mus musculu
35	39	42.4	111	1 VPX_HV2KR	O74122 human immun
36	39	42.4	112	1 VPX_HV2CA	P24110 human immun
37	39	42.4	112	1 VPX_HV2G1	P18045 human immun
38	39	42.4	112	1 VPX_HV2ST	P20881 human immun
39	39	42.4	260	1 VSP1_AGRHP	O9Y912 agkistrodon
40	39	42.4	313	1 YDJB_SCHPO	P87059 schizosacch
41	39	42.4	332	1 CX44_HUMAN	P35212 homo sapien
42	39	42.4	373	1 TRMU_STRPY	P58075 streptococc
43	39	42.4	423	1 TBB2_PORPU	P50260 porphyra pu
44	39	42.4	427	1 YKT5_YEAST	P36046 saccharomyc
45	39	42.4	441	1 TBB_BABBO	O04709 babesia bov

ALIGNMENTS

RESULT 1	ID	JANA_DROPS	STANDARD	PRT	146 AA.
AC	P54364	JANA_DROPS	STANDARD	PRT	146 AA.
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Sex-regulated protein janus-A.				
GN	JANA.				
OS	Drosophila pseudoobscura (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;				
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
ON	NCBI_TaxId=7237;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=95214615; PubMed=7700229;				
RA	Yanicosas C., Ferrer P., Vincent A., Lepesant J.-A.;				
RT	* Separate cis-regulatory sequences control expression of serendipity				
RT	beta and janus A, two immediately adjacent Drosophila genes.*;				
RL	Mol. Gen. Genet. 246:549-560(1995).				
CC	1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-B.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb.ch/announce/				
CC	or send an email to license@isb.ch).				
DR	EMBL: S77099; AAB33912.2; -				
DR	FLYBASE: FBgn0015151; Dpse\Jana.				
DR	SEQUENCE 146 AA; 16291 MW; 02F7DRC570070D7B CRC64;				
QY	Query Match	64.1%;	Score 59;	DB 1;	Length 146;
QY	Best Local Similarity	66.7%;	Pred. No. 0.0052;		
DB	Matches 10;	Conservative 1;	Mismatches 4;	Indels 0;	Gaps 0;
DB	1 DCECLGGGRSHSQSD 15				
DB	91 DTECLGGGRSHSQSD 105				
RESULT 2	JANA_DROME	STANDARD	PRT	135 AA.	
ID	JANA_DROME	STANDARD	PRT	135 AA.	
AC	P20348; Q9VAB6;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Sex-regulated protein janus-A.				
GN	JANA OR CG7933.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;				
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				

OC Muscomorpha: Ephydroidea: Drosophilidae: Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canon-S;
 RX MEDLINE=89343970; PubMed=2503707;
 RA Yancopoulos C., Vincent A., Lepesant J.-A.;
 RT "Transcriptional and posttranscriptional regulation contributes to
 RT the sex-regulated expression of two sequence-related genes at the
 RT Janus locus of *Drosophila melanogaster*.";
 RL Mol. Cell. Biol. 9:2526-2535(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-O., Andrews-Prankoch C., Baldwin D.,
 RA Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Hock J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegan C.,
 RA Jaitai M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Meklov G., Mishina N.V., Moharry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weissstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- TISSUE SPECIFICITY: SOMATIC AND GERM LINE CELLS.
 CC -1- DEVELOPMENTAL STAGE: THE NON-SEX-SPECIFIC TRANSCRIPT IS PRESENT AT
 CC ALL STAGES.
 CC -1- MISCELLANEOUS: TRANSCRIPTION OF JANA GIVES RISE TO TWO
 CC SEX-SPECIFIC AND TO ONE NON-SEX-SPECIFIC TRANSCRIPTS.
 CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-B.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M27033; AAC34203.1; -
 CC EMBL: M27033; AAC34202.1; ALT INIT.
 CC EMBL: AE003772; AAP56997.1; -
 CC PIR: A32317; A32317
 CC Flybase: FBgn0001280; jana.

SO SEQUENCE 135 AA; 15220 MW; 2720237CE7FE3132 CRC64;
 Query Match 51.1%; Score 47; DB 1; Length 135;
 Best Local Similarity 61.5%; Pred. No. 0.5;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 3 ECLGGGRISHQSO 15
 11111111
 DB 82 ECPGGGRLEHNP 94
 RESULT 3
 SYN_BRUMA
 ID SYN_BRUMA STANDARD; PRT; 548 AA.
 AC P10723;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Asparaginyl-tRNA synthetase, cytoplasmic (BC 6.1.1.22) (Asparagine--
 DE tRNA ligase) (AsnRS) (Potentially protective 63 kDa antigen).
 OS Brugia malayi.
 OC Eukaryota: Metazoa: Nematoda: Chromadorea: Spirurida: Filarioidae;
 OC Onchocercidae: Brugia.
 OX NCBI_TaxID=6279;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8828341; PubMed=2840577;
 RA Perrine K.G., Denker J.A., Nilsen T.W.;
 RT "A multi-copy gene encodes a potentially protective antigen in *Brugia*
 RT malayi.";
 RL Mol. Biochem. Parasitol. 30:97-104(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88217950; PubMed=3368467;
 RA Nilsen T.W., Maroney P.A., Goodwin R.G., Perrine K.G., Denker J.A.,
 RA Nandori J., Kazura J.W.;
 RT "Cloning and characterization of a potentially protective antigen in
 RT lymphatic filariasis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:3604-3607(1988).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=96049569; PubMed=7589498;
 RA Krom M., Marguard K., Hartlein M., Price S., Leberman R.;
 RT "An immunodominant antigen of *Brugia malayi* is an asparaginyl-tRNA
 RT synthetase.";
 RL FEBS Lett. 374:122-124(1995).
 CC -1- FUNCTION: POTENTIALLY PROTECTIVE ANTIGEN IN LYMPHATIC FILARIASIS.
 CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
 CC diphosphate + L-asparaginyl-tRNA(Asn).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J03971; AAA27852.1; -
 CC EMBL: J03266; AAA27849.1; -
 CC PIR: A28209; A28209.
 CC PIR: A54510; A54510.
 CC InterPro: IPR002106; AAIRNA_ligaseII.
 CC InterPro: IPR004522; Asns.
 CC InterPro: IPR004364; tRNA-synt.2.
 CC InterPro: IPR004365; tRNA-anti.
 CC Pfam: PF00152; tRNA-synt.2; 1.
 CC Pfam: PF01336; tRNA-anti.1;
 CC TIGRfam: TIGR00457; asns; 1.
 CC PROSITE: PS50862; AA-TRNA_LIGASE_II; 1.
 CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

```

RESULT 4
Y4HP_RHISN
ID Y4HP_RHISN STANDARD: PRT: 552 AA.
AC P50360.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 61.7 kDa protein Y4HP.
GN Y4HP.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Frielberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and Legumes.";
RL Nature 387:394-401(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX Rochepeau P., Fellay R., Broughton W.J.;
RL Submitted (JUL-1993) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO A NUMBER OF ORF OR VARIABLE SIZE IN R.MELLIOTI AND
CC A.TUMEFACIENS TI PLASMID.
CC -1- SIMILARITY: TO Y4AO, Y4UD AND Y4OI.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO TWO
CC FRAMESHIFTS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74068; CAA52197.1; -
DR EMBL: AE000077; AAB92454.1; -
DR InterPro: IPR004291; Transposase_25.
DR Pfam: PF03050; Transposase_25.1.
KW Hypothetical protein; Plasmid.
FT CONFLICT 62 62 R -> S (IN REF. 2).
FT CONFLICT 104 104 R -> S (IN REF. 2).
FT CONFLICT 118 118 R -> G (IN REF. 2).
SQ SEQUENCE 552 AA: 61/747 MW: 9A2A3258AB9F7CDD CRC64;

Query Match 47.8%; Score 44; DB 1; Length 552;
Best Local Similarity 46.7%; Pred. No. 6.9;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
2 CECGAGGRISHOSOD 16
| | | | | | | | | |

```

```

Db      138 CTCGGGSRLLKGED 152

RESULT 5
MTHS_YEAST
ID      MTHS_YEAST      STANDARD;      PRT;      599 AA.
AC      P53128; 092318;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Methylenetetrahydrofolate reductase 2 (EC 1.5.1.20) .
GN      MET13 OR MET11 OR YGL125W OR G2882.
OS      Saccharomyces cerevisiae (Baker's Yeast) .
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX      NCBI_TaxID=4932;
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97051592; PubMed=8896269;
RA      Tizon B., Rodriguez-Torres A.M., Rodriguez-Belmonte E., Cadahia J.L.,
RA      Cerdan E.;
RT      *Identification of a putative methylenetetrahydrofolate reductase by
RT      sequence analysis of a 6.8 kb DNA fragment of yeast chromosome VII.*;
RL      Yeast 12:1047-1051(1996).
CC      -1 CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-
CC      methylenetetrahydrofolate + NADPH.
CC      -1 COFACTOR: FAD (BY SIMILARITY).
CC      -1 PATHWAY: Folate metabolism.
CC      -1 SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC      (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; 272647; CAA96833.1; -.
DR      EMBL; X94106; CAA63833.1; -.
DR      HSSP; P00394; 1B5T.
DR      SGD; S0003093; MET13.
DR      InterPro; IPR004621; Fadh2_euk.
DR      InterPro; IPR003171; mehydrof_redctse.
DR      Pfam; PF02219; MTHFR; 1.
DR      TIGRFAMs; TIGR00677; fadh2_euk; 1.
KW      Oxidoreductase; Flavoprotein; FAD; NADP.
FT      CONFLECT 73
FT      73 A -> R (IN REF. 1; CAA63833) .
SQ      SEQUENCE 599 AA; 68472 MW; AD3455BB52A4E700 CRC64;

Query Match      47.8%; Score 44; DB 1; Length 599;
Best Local Similarity 72.7%; Pred. No. 7.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      6 GGGRIHQSQD 16
      |||||
Db      58 GGGRIHSLSTD 68

RESULT 6
TBBL_TRIVI
ID      TBBL_TRIVI      STANDARD;      PRT;      446 AA.
AC      P31864;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      Tubulin beta-1 chain.
DE      TUB1.
OS      Trichoderma viride.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Hypocreales; mitosporic Hypocreales; Trichoderma.
NCBI_TaxID=5547;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-79 BR47;
 RA MEDLINE-93341462; PubMed-8341264;
 RA Goldman G.H., Temmerman W., Herrera-Estrella A., Jacobs D.,
 RA Contreras R., van Montagu M.;
 RT "A nucleotide substitution in one of the beta-tubulin genes of
 RT Trichoderma viride confers resistance to the antimitotic drug methyl
 RT benzimidazole-2-yl-carbamate.";
 RL Mol. Gen. Genet. 240:73-80(1993).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 215054; CAA78764.1; -
 CC PIR; S25553; S25553.
 CC InterPro; IPR002453; Beta_tubulin.
 CC InterPro; IPR000217; Tubulin.
 CC InterPro; IPR003008; Tubulin_FtsZ.
 CC Pfam; PF00091; tubulin; 1.
 CC PRINTS; PR01161; TUBULIN.
 CC PROSITE; PS00227; TUBULIN; 1.
 CC PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
 CC Microtubules; GTP-binding.
 CC NP_BIND 142 148 GTP (POTENTIAL).
 CC SEQUENCE 446 AA; 49625 MW; B43E7B478B1C13 CRC64;
 SQ
 Query Match 46.7%; Score 43; DB 1; Length 446;
 Best Local Similarity 58.3%; Pred. No. 8.1;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DCCRLGGGRISH 12
 DB 128 NCECLGFGFITH 139
 ID EXON_HCMVA STANDARD; PRT; 584 AA.
 AC P16789;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Alkaline exonuclease (EC 3.1.11.-).
 OS U98.
 GN Human cytomegalovirus (strain AD169).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_Taxid=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90269039; PubMed-2161319;
 RA Choe M.S., Bankier A.T., Beck S., Bohnl R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrett B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC [2]
 CC SEQUENCE OF 465-584 FROM N.A.
 CC MEDLINE-8906836; PubMed-2535729;
 CC Martinez J., Lahiani R.S., St Jeor S.C.;
 CC "Analysis of a region of the human cytomegalovirus (AD169) genome

RT coding for a 25-kilodalton virion protein.";
 RL J. Virol. 63:233-241(1989).
 RN [3]
 RP SEQUENCE OF 465-584 FROM N.A.
 RX MEDLINE-91087306; PubMed-1845897;
 RA Lahiani R.S., Oateson E.W., Adlisch J.D., St Jeor S.C.;
 RT "Characterization of a human cytomegalovirus 1.6-kilobase late mRNA
 RT and identification of its putative protein product.";
 RL J. Virol. 65:373-381(1991).
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES ALKALINE EXONUCLEASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X17403; CAA35334.1; -
 CC EMBL; M62615; AAA45905.1; -
 CC PIR; S09863; COBEJ6.
 CC InterPro; IPR001616; Herpes_alk_exo.
 CC Pfam; PF01771; Herpes_alk_exo; 1.
 CC PRINTS; PR00924; ALKEXNUCLASE.
 CC Hydrolase; Nuclease; Exonuclease.
 CC SEQUENCE 584 AA; 65272 MW; C881D722B8014126 CRC64;
 SQ
 Query Match 46.7%; Score 43; DB 1; Length 584;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 CECILGGGRISH 12
 DB 511 CELLAGGRVPH 521
 ID FLHC_ECOLI STANDARD; PRT; 192 AA.
 AC P11165; P76303;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flagellar transcriptional activator FlhC.
 GN FLHC OR FLAI OR B1891 OR Z2945 OR ECS2601.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88169478; PubMed-2832369;
 RA Bartlett D.H., Frantz B.B., Matsumura P.;
 RT "Flagellar transcriptional activators FlhB and FlhA: gene sequences
 RT and 5' consensus sequences of operators under FlhB and FlhA control.";
 RL J. Bacteriol. 170:1575-1581(1988).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN-K12 / MG1655;
 CC MEDLINE-97426617; PubMed-9278503;
 CC Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 CC Blythe M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 CC Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 CC Mau B., Shao Y.;
 CC "The complete genome sequence of Escherichia coli K-12.";
 CC Science 277:1453-1474(1997).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN-K12;
 CC MEDLINE-97251358; PubMed-9097040;

RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
 RA Saito N., Sampei G., Seki Y., Sivasubramanian S., Tagami H.,
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horuchi T.,
 RT "A 460-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 40,150,000 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / ATCC 700927;
 RX MEDLINE-21074935; PubMed-11206551;
 RA Berta N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirpatrick H.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamianos K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RX MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 [6]
 RP FUNCTION: Transcriptional activator. Together with flhD it acts as
 CC a compound sigma factor that activates class 2 flagellar genes.
 CC -1- SIMILARITY: BELONGS TO THE FLHC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M19439; AAA23788.1; -
 DR EMBL: AF000283; AAC74961.1; -
 DR EMBL: D90831; BAA15712.1; -
 DR EMBL: AE005411; AAG56881.1; -
 DR EMBL: AP002559; BAB36024.1; -
 DR PIR: B27735; XMECIF.
 DR Ecogene: EG10319; flhC.
 KW Transcription: Transcription regulation: Sigma factor;
 KW DNA-directed RNA polymerase; DNA-binding; Flagella; Complete proteome.
 FT CONFLICT 149 149 H -> D (IN REF. 1).
 SQ SEQUENCE 192 AA; 21566 MW; 4504AF0580545C0C CRC64;
 Query Match 45.7%; Score 42; DB 1; Length 192;
 Best Local Similarity 50.0%; Pred. No. 5;
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 2 CECIAGGRISHQSQ 15
 DB 137 CCGCGGNETFHAHQ 150
 ID FLHC_SALTY STANDARD; PRT; 192 AA.
 AC 052222;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flagellar transcriptional activator flhC.

GN FLHC OR STM1924.
 OS *Salmonella typhimurium*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Salmonella*.
 OX NCBI_TaxID=602;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2;
 RA Touguchi A., Harshey R.M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20052966; PubMed-10586519;
 RA Yanagihara S., Iyoda S., Onishi K., Ino T., Kutukake K.,
 RT "Structure and transcriptional control of the flagellar master operon
 RT of *Salmonella typhimurium*.";
 RL Genes genet. Syst. 74:105-111(1999).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:856-856(2001).
 [4]
 RP FUNCTION: Transcriptional activator. Together with flhD it acts as
 CC a compound sigma factor that activates class 2 flagellar genes (by
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE FLHC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF029300; AAB96440.1; -
 DR EMBL: D43640; BAA85315.1; -
 DR EMBL: AE008786; AAL20340.1; ALT_INIT.
 DR StyGene: SG10643; flhC.
 KW Transcription: Transcription regulation: Sigma factor;
 KW DNA-directed RNA polymerase; DNA-binding; Flagella; Complete proteome.
 SQ SEQUENCE 192 AA; 21579 MW; C734C691ECC3F8D CRC64;
 Query Match 45.7%; Score 42; DB 1; Length 192;
 Best Local Similarity 50.0%; Pred. No. 5;
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 2 CECIAGGRISHQSQ 15
 DB 137 CCGCGGNETFHAHQ 150
 ID FLHC_SERMA STANDARD; PRT; 194 AA.
 AC 085807;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flagellar transcriptional activator flhC.
 OS *Serratia marcescens*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Serratia*.
 OX NCBI_TaxID=615;
 RN (1)

```

RP SEQUENCE FROM N.A.
RC STRAIN-CH-1:
RA Liu J.H., Lai M.T., Ang S., Shu J.C., Soo P.C., Hong Y.T., Yi W.C.,
RA Lai H.C., Luh K.T., Ho S.W., Swift S.;
RT "Coordinate regulation of nuc. cell division and flagellar synthesis
RT by flhdc in Serratia marcescens.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Transcriptional activator. Together with flhnd it acts as
CC a compound sigma factor that activates class 2 flagellar genes (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE FLHC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF077334; AAC27634.1;
KW Transcription: Transcription regulation; Sigma factor;
KW DNA-directed RNA polymerase; DNA-binding; Flagella.
SQ SEQUENCE 194 AA; 21809 MW; E57993DBZ73AD2DD CRC64;

Query Match 45.7%; Score 42; DB 1; Length 194;
Best Local Similarity 50.0%; Pred. No. 5.1;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 CECLGGGRSHSQ 15
Db 139 CSCGGAFTHAQ 152

RESULT 11
FLHC_XENNE
ID FLHC_XENNE STANDARD: PRT; 194 AA.
AC Q9X9F3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flagellar transcriptional activator flhc.
GN FLHC.
OS Xenorhabdus nematophilus.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Xenorhabdus.
OX NCBI_TaxId=628;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FL;
RA Givaudan A.G., Lanois A.;
RT "flhdc gene disruptions leads to pleiotropic phenotypes.";
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Transcriptional activator. Together with flhnd it acts as
CC a compound sigma factor that activates class 2 flagellar genes (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE FLHC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF012828; CAB41408.1;
KW Transcription: Transcription regulation; Sigma factor;
KW DNA-directed RNA polymerase; DNA-binding; Flagella.
SQ SEQUENCE 194 AA; 21852 MW; 4EC38ADD00908677 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 194;
Best Local Similarity 46.7%; Pred. No. 5.1;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

```

OY      1  DOECLGGGRISHQSO 15
      :|:|:|:|:|:|
DB      138  BCRCCGCFITTHAQ 152

RESULT 12
TRMU_BACHD
ID      TRMU_BACHD      STANDARD:      PRT:      371 AA.
AC      O9KDF2:
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
      (EC 2.1.1.61).
GN      TRMU OR BH1261.
OS      Bacillus halodurans.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=86665;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C-125 / JCM 9153;
RX      MEDLINE=20512582; PubMed=11058132;
RA      Takami H., Nakasone C., Takaki Y., Maeno G., Sasaki R., Masui N.,
      Fujii F., Hirama O., Nakamura Y., Ogasawara N., Kunara S.,
      Horikoshi K.;
RT      "Complete genome sequence of the alkaliphilic bacterium Bacillus
      halodurans and genomic sequence comparison with Bacillus subtilis.";
RL      Nucleic Acids Res. 28:4317-4331(2000).
CC      -1- CATABOLIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
      homocysteine + tRNA containing 5-methylaminomethyl-2-
      thiouridylate.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC      -1- SIMILARITY: BELONGS TO THE TRMU FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
      the European Bioinformatics Institute. There are no restrictions on its
      use by non-profit institutions as long as its content is in no way
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (see http://www.isb-sdb.ch/announce/
      or send an email to license@sdb-sib.ch).
CC      -----
DR      EMBL; AP001511; BAB04980.1; -.
DR      InterPro; IPR004506; Trmu.
DR      InterPro; IPR004135; tRNA_Me_trans.
DR      Pfam; PF03054; tRNA_Me_trans; 1.
DR      TIGRFAMs; TIGR00420; trmu; 1.
SQ      Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ      SEQUENCE 371 AA; 41625 MW; ECIPEA4DA92DB24 CRC64;

Query Match      45.7%; Score 42; DB 1; Length 371;
Best Local Similarity 77.8%; Pred. No. 9.9;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      4  CUCGGGRISH 12
      |||||
DB      352  CUCGGGTIDH 360

RESULT 13
C1BB_BACTU
ID      C1BB_BACTU      STANDARD:      PRT:      1229 AA.
AC      045739:
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Pesticialial crystal protein cryIbB (insecticidal delta-endotoxin
      de cryIb(b)) (crystalline entomocidal protoxin) (140 kDa crystal protein).
GN      CRYIIB OR CRYIIB(B) OR CRYET5.
OS      Bacillus thuringiensis.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

```

OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STEAIN=E55847 / NRRL B-21110;
 RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
 RT "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
 RT toxic to lepidopteran insects."
 RL Patent number US5322687, 21-JUN-1994.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L32020; AAA2344.1; -
 DR HSSP: P02965; ICIT.
 DR InterPro: IPR001178; Endotoxin.
 DR Pfam: PF00555; endotoxin.1.
 DR Toxin; Sporulation.
 SK SEQUENCE 1229 AA; 139769 MW; AAC949DB675C3269 CRC64;
 OY 1 DECELCGGRISHOS 14
 Db 863 DCSCRGRCRHHS 876
 Query Match 45.7%; Score 42; DB 1; Length 1229;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

RESULT 14
 ID C1BD_BACTZ STANDARD; PRT; 1231 AA.
 AC 092A25;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pesticidal crystal protein cryIbd (insecticidal delta-endotoxin
 DE CRYIB(d)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
 GN CRYIBD OR CRYIB(D) OR CRYIA1 OR CRYEL.
 OS Bacillus thuringiensis (subsp. wuhanensis).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=52024;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-525;
 RA MEDLINE=20153386; PubMed=1068690;
 RA Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
 RT "Cloning of two new cry genes from Bacillus thuringiensis subsp.
 RT wuhanensis strain."
 RL Curr. Microbiol. 40:227-232(2000).
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF LEPIDOPTERAN LARVAE. TOXIC TO PLOUTELLA
 CC XYLOSTELLA.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U70726; AAD10292.1; -
 DR HSSP: P02965; ICIT.
 DR InterPro: IPR001178; Endotoxin.
 DR Pfam: PF00555; endotoxin.1.
 DR Toxin; Sporulation.
 SK SEQUENCE 1231 AA; 139654 MW; 129A0371CDDBE52 CRC64;
 OY 1 DECELCGGRISHOS 14
 Db 865 DCSCRGRCRHHS 878
 Query Match 45.7%; Score 42; DB 1; Length 1231;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

RESULT 15
 ID F1NC_BOVIN STANDARD; PRT; 2265 AA.
 AC P07589;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fibronectin (FN).
 GN FNI.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=87054047; PubMed=3780752;
 RA Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
 RT "Complete primary structure of bovine plasma fibronectin."
 RL Eur. J. Biochem. 161:441-453(1986).
 RN [2]
 RP PARTIAL SEQUENCE.
 RA MEDLINE=83117805; PubMed=6218503;
 RA Petersen T.E., Thorgeresen H.C., Skorstengaard K., Vibe-Pedersen K.,
 RA Sahl P., Solttrup-Jensen L., Magnusson S.;
 RT "Partial primary structure of bovine plasma fibronectin: three types
 RT of internal homology."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).
 RN [3]
 RP SEQUENCE OF 2170-2265 FROM N.A.
 RA MEDLINE=83221567; PubMed=6304699;
 RA Kornblith A.R., Vibe-Pedersen K., Baralle F.E.;
 RT "Isolation and characterization of cDNA clones for human and bovine
 RT fibronectins."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).
 CC -1- FUNCTION: FIBRONECTIN BIND CELL SURFACES AND VARIOUS COMPOUNDS
 CC INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
 CC ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
 CC HEALING, AND MAINTENANCE OF CELL SHAPE.
 CC -1- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
 CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
 CC TO A LESSER EXTENT HOMODIMERS.
 CC -1- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
 CC STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
 CC OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
 CC BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
 CC FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
 CC DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.
 CC -1- P.M.: SULFATED (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.
 CC -1- SIMILARITY: CONTAINS 15 FIBRONECTIN TYPE III DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: K00800; AAA30521.2; -
 DR PIR: A26452; FNBO.
 DR HSSP: P02751; 2FN2.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR000562; FN_Type_II.
 DR InterPro: IPR000083; Fibrinctn.
 DR InterPro: IPR003962; FnIII_repeat.
 DR Pfam: PF00039; fnI; 12.
 DR Pfam: PF00040; fn2; 2.
 DR Pfam: PF00041; fn3; 15.
 DR PRINTS: PRO0012; FNTYPEI.
 DR PRINTS: PRO0013; FNTYPEII.
 DR PRINTS: PRO0014; FNTYPEIII.
 DR ProDom: PD000995; FN_Type_II; 2.
 DR SMART: SM00058; FN1; 12.
 DR SMART: SM00059; FN2; 2.
 DR SMART: SM00060; FN3; 14.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS00023; FIBRONECTIN_2; 2.
 DR PROSITE: PS01253; FIBRONECTIN_1; 12.
 DR PROSITE: PS01253; FIBRONECTIN_1; 12.
 KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;
 KW Sulfation; Cell adhesion; Repeat; Alternative splicing.
 FT MOD_RES 1 241 1
 FT DOMAIN 277 577
 FT DUA_BIND 876 1141
 FT DOMAIN 1236 1509
 FT DOMAIN 1600 1870
 FT DOMAIN 1991 2216
 FT DOMAIN 19 59
 FT DOMAIN 64 107
 FT DOMAIN 108 151
 FT DOMAIN 153 197
 FT DOMAIN 198 242
 FT DOMAIN 275 314
 FT DOMAIN 314 373
 FT DOMAIN 374 438
 FT DOMAIN 437 480
 FT DOMAIN 485 527
 FT DOMAIN 528 571
 FT DOMAIN 578 669
 FT DOMAIN 688 778
 FT DOMAIN 779 874
 FT DOMAIN 875 964
 FT DOMAIN 965 1054
 FT DOMAIN 1055 1141
 FT DOMAIN 1142 1234
 FT DOMAIN 1235 1325
 FT DOMAIN 1326 1415
 FT DOMAIN 1416 1509
 FT DOMAIN 1510 1599
 FT DOMAIN 1600 1691
 FT DOMAIN 1692 1780
 FT DOMAIN 1781 1870
 FT DOMAIN 1871 1990
 FT DOMAIN 1982 2061
 FT DOMAIN 2083 2127
 FT DOMAIN 2128 2170
 FT DOMAIN 2172 2215
 FT SITE 1493 1495

FT DISULFID 21 47
 FT DISULFID 45 56
 FT DISULFID 66 94
 FT DISULFID 92 104
 FT DISULFID 110 138
 FT DISULFID 136 148
 FT DISULFID 155 184
 FT DISULFID 182 194
 FT DISULFID 200 229
 FT DISULFID 227 239
 FT DISULFID 277 304
 FT DISULFID 302 311
 FT DISULFID 329 355
 FT DISULFID 343 370
 FT DISULFID 389 415
 FT DISULFID 403 430
 FT DISULFID 439 467
 FT DISULFID 465 477
 FT DISULFID 487 514
 FT DISULFID 512 524
 FT DISULFID 530 558
 FT DISULFID 556 568
 FT DISULFID 2085 2114
 FT DISULFID 2112 2124
 FT DISULFID 2130 2157
 FT DISULFID 2155 2167
 FT DISULFID 2174 2200
 FT DISULFID 2198 2209
 FT DISULFID 2246 2246
 FT DISULFID 2250 2250
 FT MOD_RES 845 845
 FT MOD_RES 850 850
 FT CARBOHYD 399 399
 FT CARBOHYD 497 497
 FT CARBOHYD 511 511
 FT CARBOHYD 846 846
 FT CARBOHYD 976 976
 FT CARBOHYD 1213 1213
 FT CARBOHYD 1987 1987
 FT CARBOHYD 1943 1943
 FT CARBOHYD 1944 1944
 FT MOD_RES 2263 2263
 SQ SEQUENCE 2265 AA; 249557 MW; C2D21D486F498D5C CRC64;
 Query Match 45.7%; Score 42; DB 1; Length 2265;
 Best Local Similarity 66.7%; Pred. No. 65;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DCECLGGR 9
 II I: I I I
 Db 91 DCTCIGAGR 99
 Search completed: May 7, 2003, 19:03:58
 Job time : 3.90826 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:00:25 ; Search time 11.0826 Seconds
(without alignments)
297.472 Million cell updates/sec

Title: US-09-914-831-3

Perfect score: 92

Sequence: 1 DCECLGGGRISHQSD 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Match	Length	DB	ID	Description
1	92	100.0	124	11	Q9DAK9	Q9DAK9 mus musculi
2	92	100.0	125	4	Q9NRX4	Q9NRX4 homo sapien
3	87	94.6	125	4	Q9H0Y3	Q9H0Y3 homo sapien
4	49	53.3	244	8	Q8W9R2	Q8W9R2 mesostigma
5	47	51.1	135	5	Q9BHA8	Q9BHA8 drosophila
6	46	50.0	102	5	Q9BH68	Q9BH68 drosophila
7	46	50.0	102	5	Q9BH67	Q9BH67 drosophila
8	46	50.0	104	5	Q9BM25	Q9BM25 drosophila
9	46	50.0	104	5	Q9BM21	Q9BM21 drosophila
10	46	50.0	109	5	Q9BM27	Q9BM27 drosophila
11	46	50.0	115	5	P90861	P90861 caenorhabdi
12	46	50.0	119	5	Q8T475	Q8T475 drosophila
13	46	50.0	135	5	Q9BM98	Q9BM98 drosophila
14	46	50.0	135	5	Q9BM97	Q9BM97 drosophila
15	46	50.0	135	5	Q9BM96	Q9BM96 drosophila
16	46	50.0	135	5	Q9BM99	Q9BM99 drosophila

17	46	50.0	135	5	Q95VY7	Q95VY7 drosophila
18	46	50.0	758	4	Q9C0B5	Q9C0B5 homo sapien
19	45	48.9	394	16	Q9KCP6	Q9KCP6 bacillus ha
20	45	48.9	402	2	Q31267	Q31267 agrobacteri
21	44	47.8	184	2	Q9X984	Q9X984 rhizobium m
22	44	47.8	648	5	Q9NKD7	Q9NKD7 drosophila
23	44	47.8	701	5	Q9VJU4	Q9VJU4 drosophila
24	44	47.8	1282	4	Q8TER0	Q8TER0 homo sapien
25	43	46.7	117	11	Q9D596	Q9D596 mus musculi
26	43	46.7	373	16	Q9T2G7	Q9T2G7 streptococc
27	43	46.7	477	5	Q9T2G7	Q9T2G7 caenorhabdi
28	43	46.7	584	12	Q99D32	Q99D32 human cytom
29	42	45.7	164	2	Q08242	Q08242 pseudomonas
30	42	45.7	192	2	Q8VND3	Q8VND3 salmonella
31	42	45.7	194	16	Q8Z505	Q8Z505 salmonella
32	42	45.7	269	10	Q9C5W1	Q9C5W1 arabidopsis
33	42	45.7	367	10	Q9LDP9	Q9LDP9 arabidopsis
34	42	45.7	379	4	Q95610	Q95610 homo sapien
35	42	45.7	387	5	Q9BKC8	Q9BKC8 jakoba inca
36	42	45.7	387	8	Q9ZTU3	Q9ZTU3 guillardia
37	42	45.7	387	10	Q9ZTU5	Q9ZTU5 guillardia
38	42	45.7	387	10	Q9SQJ5	Q9SQJ5 guillardia
39	42	45.7	441	8	Q9SEV2	Q9SEV2 guillardia
40	42	45.7	567	4	Q8WUL3	Q8WUL3 homo sapien
41	42	45.7	642	4	Q9H1B8	Q9H1B8 homo sapien
42	42	45.7	734	8	Q9THV8	Q9THV8 rondeletia
43	42	45.7	1043	10	Q9FNC3	Q9FNC3 arabidopsis
44	42	45.7	1061	13	Q919K0	Q919K0 xenopus lae
45	42	45.7	1140	4	Q96KG7	Q96KG7 homo sapien

ALIGNMENTS

RESULT 1
ID Q9DAK9 PRELIMINARY; PRT; 124 AA.
AC Q9DAK9;
DT 01-JUN-2001 (TREMBLrel, 17, Created)
DT 01-JUN-2001 (TREMBLrel, 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel, 21, Last annotation update)
DE 1700008C22Rik Protein (RIKEN CDNA 1700008C22 gene).
GN 1700008C22Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP
[1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giusti C., King B., Kochiwa H.,
RA Kuehl L.M., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Hoffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Kyrushaw-Boriss A., Yoshida K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN
RP SEQUENCE FROM N.A.

RC TISSUE-MAMMARY GLAND;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK005756; BAB24222.1; -
 DR EMBL; BC028657; AAH28657.1; -
 DR MGD; MGI:1922704; 170008C22Rik.
 SQ SEQUENCE 124 AA; 13996 MW; 583F9A3CE95A2F66 CRC64;

Query Match 100.0%; Score 92; DB 11; Length 124;
 Best Local Similarity 100.0%; Pred. No. 2,3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DCECLGGGRISHOSQ 16
 DB 69 DCECLGGGRISHOSQ 84

RESULT 2

O9NRX4 PRELIMINARY; PRT; 125 AA.
 AC O9NRX4;
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Sex-regulated protein Janus-A (CGI-202) (Similar to HSPC141 protein).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ADRENAL GLAND;
 RA Li Y., Shi J., Huang C., Ren S., Zhou J., Yu Y., Xu S.,
 Wang Y., Fu G., Chen Z., Han Z.;
 RT "A novel gene expressed in human adrenal gland."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ADRENAL GLAND;
 RA Peng Y., Gu Y., Li Y., Fu S., Gu J., Gu W., Jiang C., Yu Y., Han Z.,
 Wang Y., Chen Z., Fu G.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lin W.-C.;
 RT "Identification of novel human genes by comparative proteomics."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Strausberg R.;
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF164795; AAP80759.1; -
 DR EMBL; AF285119; AAG01156.1; -
 DR EMBL; BC024648; AAH24648.1; -
 SQ SEQUENCE 125 AA; 13832 MW; 24F0CA2BADB78478 CRC64;

Query Match 100.0%; Score 92; DB 4; Length 125;
 Best Local Similarity 100.0%; Pred. No. 2,3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSQ 16
 DB 70 DCECLGGGRISHOSQ 85

RESULT 3
 O9HOY3 PRELIMINARY; PRT; 125 AA.
 AC O9HOY3;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 13.8 kDa protein.
 GN DKFZ564M173.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA MEDLINE-21154917; PubMed-11230166;
 RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,
 Ansoorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
 Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 Mewes H.W., Ottenwaelder B., Obermayer B., Tampe J., Heubner D.,
 Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs."
 RL Genome Res. 11:422-435(2001).
 DR EMBL; AL136644; CAB6579.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 125 AA; 13822 MW; 3C475F5B43ADED5A CRC64;

Query Match 94.6%; Score 87; DB 4; Length 125;
 Best Local Similarity 93.8%; Pred. No. 1,3e-06;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DCECLGGGRISHOSQ 16
 DB 70 DCECLGGGRISHOSQ 85

RESULT 4

O8W9R2 PRELIMINARY; PRT; 244 AA.
 AC O8W9R2;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical intron-encoded protein.
 GN ORF244.
 OS Mesostigma viride.
 OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatozoa;
 OC Mesostigmatales; Mesostigmataceae; Mesostigma.
 OX NCBI_TaxID=41882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIES-296;
 RA MEDLINE-21624446; PubMed-11752187;
 RA Turmel M., Odis C., Lemieux C.;
 RT "The Complete Mitochondrial DNA Sequence of Mesostigma viride
 RT Identifies This Green Alga as the Earliest Green Plant Divergence and
 RT Predicts a Highly Compact Mitochondrial Genome in the Ancestor of All
 RT Green Plants."
 RL Mol. Biol. Evol. 19:24-38(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIES-296;
 RA Turmel M., Odis C., Lemieux C.;
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF353999; AAL36746.1; -
 DR InterPro; IPR001982; Endomuc_LAG/HNH.
 DR Pfam; PF00961; LAGL1DAG_1; 2.
 KW Mitochondrion.
 SQ SEQUENCE 244 AA; 28753 MW; 7345C045472FF08 CRC64;

Query Match 53.3%; Score 49; DB 8; Length 244;
 Best Local Similarity 57.1%; Pred. No. 4;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 CECIGGGGRISHOSQ 15

Db 51 CQCLMGGRITNSSK 64

RESULT 5

09BH48 PRELIMINARY; PRT: 135 AA.

AC 09BH48; 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Janus A.
 GN JANA.
 OS Drosophila telesiieri (Fruit fly), and
 OS Drosophila yakuba (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7243, 7245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-D.telesiieri, and D.yakuba;
 RA Parsch J., Melkleyohn C.D., Hartl D.L.;
 RT "Molecular evolution of the ocnus and janus genes in the Drosophila
 melanogaster species subgroup.";
 RT Submitted (Nov-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF013342; AAG50363.1; -;
 DR EMBL: AF013341; AAG50362.1; -;
 DR FlyBase: FBgn0043602; Dyak\Jana.
 DR FlyBase: FBgn0043616; Dtel\Jana.
 SQ SEQUENCE 135 AA; 15206 MW; C1756A84DE05671C CRC64;

Query Match 51.1%; Score 47; DB 5; Length 135;
 Best Local Similarity 61.5%; Pred. No. 4.6;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 ECLGGGRISHSQ 15
 Db 82 ECPGGGRIEHNP 94

RESULT 6
 09BH68 PRELIMINARY; PRT: 102 AA.

AC 09BH68; 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KENYA.5 AND KENYA.12;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 Drosophila simulans Complex Species.";
 RT Genetics 156:1913-1931(2000).
 RL EMBL: AF284454; AAG49468.1; -;
 DR EMBL: AF284455; AAG49470.1; -;
 DR FlyBase: FBgn0016348; Dsim\Jana.
 FT NON_TER 1
 SQ SEQUENCE 102 AA; 11608 MW; 7C1AC80B7E9369EB CRC64;

Query Match 50.0%; Score 46; DB 5; Length 102;
 Best Local Similarity 80.0%; Pred. No. 5;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 Db 49 ECPGGGRIEH 58

RESULT 7

09BH67 PRELIMINARY; PRT: 102 AA.

AC 09BH67; 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CAP MALHERREUX, AND PORT-LOUIS;
 RA MEDLINE-20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 Drosophila simulans Complex Species.";
 RT Genetics 156:1913-1931(2000).
 RL EMBL: AF284457; AAG49474.1; -;
 DR EMBL: AF284458; AAG49476.1; -;
 DR FlyBase: FBgn0043691; Dmau\Jana.
 FT NON_TER 1
 SQ SEQUENCE 102 AA; 11608 MW; 7C1AC80B7E9369EB CRC64;

Query Match 50.0%; Score 46; DB 5; Length 102;
 Best Local Similarity 80.0%; Pred. No. 5;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 Db 49 ECPGGGRIEH 58

RESULT 8
 09BM25 PRELIMINARY; PRT: 104 AA.

AC 09BM25; 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1631;
 RA MEDLINE-20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 Drosophila simulans Complex Species.";
 RT Genetics 156:1913-1931(2000).
 RL EMBL: AF284456; AAG49472.1; -;
 DR EMBL: AF284457; AAG49474.1; -;
 DR FlyBase: FBgn0043691; Dmau\Jana.
 FT NON_TER 1
 SQ SEQUENCE 104 AA; 11834 MW; 690E994E3A961437 CRC64;

Query Match 50.0%; Score 46; DB 5; Length 104;
 Best Local Similarity 80.0%; Pred. No. 5.1;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ECLGGGRISH 12
 |||||||
 Db 51 ECPGGRIEH 60

RESULT 9

Q9BMZ1 PRELIMINARY; PRT; 104 AA.

AC Q9BMZ1; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

GN JANA.

OS Drosophila sechellia (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7238;

RP SEQUENCE FROM N.A.

RC STRAIN=228;

RA MEDLINE=20556156; PubMed=1102384;

RA Kliman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,

RA Berry A.J., McCarter J., Wakeley J., Hey J.;

RT "The Population Genetics of the Origin and Divergence of the

RT Drosophila simulans Complex Species.";

RL Genetics 156:1913-1931(2000).

DR EMBL; AF284459; AAG49478.1; -;

DR FLYBase; FBgn0043637; Dsec\Jana.

FT NON_TER 1

SEQUENCE 104 AA; 11735 MW; 73BF32548B261427 CRC64;

Query Match 50.0%; Score 46; DB 5; Length 104;

Best Local Similarity 80.0%; Pred. No. 5.1;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ECLGGGRISH 12
 |||||||
 Db 51 ECPGGRIEH 60

RESULT 10

Q9BMZ7 PRELIMINARY; PRT; 109 AA.

AC Q9BMZ7; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

GN JANA.

OS Drosophila simulans (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7240;

RP SEQUENCE FROM N.A.

RC STRAIN=KENYA-2;

RA MEDLINE=20556156; PubMed=11102384;

RA Kliman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,

RA Berry A.J., McCarter J., Wakeley J., Hey J.;

RT "The Population Genetics of the Origin and Divergence of the

RT Drosophila simulans Complex Species.";

RL Genetics 156:1913-1931(2000).

DR EMBL; AF284453; AAG49466.1; -;

DR FLYBase; FBgn0016348; Dsim\Jana.

FT NON_TER 1

SEQUENCE 109 AA; 12384 MW; 1118CE3722E56A6F CRC64;

Query Match 50.0%; Score 46; DB 5; Length 109;

Best Local Similarity 80.0%; Pred. No. 5.4;

QY 3 ECLGGGRISH 12
 |||||||
 Db 56 ECPGGRIEH 65

RESULT 11

P90861 PRELIMINARY; PRT; 115 AA.

AC P90861; 01-MAY-1997 (TREMBlrel. 03, Created)

DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

GN F36A2.8 protein.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.

RA Lennard N.;

RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

RL (2)

RP SEQUENCE FROM N.A.

RA MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C. elegans: A platform for

RT investigating biology.";

RL Science 282:2012-2018(1998).

DR EMBL; Z81077; CAB03070.1; -;

FT SEQUENCE 115 AA; 13183 MW; 29BC8720EC09203 CRC64;

Query Match 50.0%; Score 46; DB 5; Length 115;

Best Local Similarity 70.0%; Pred. No. 5.7;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 ECLGGGRISH 12
 |||||||
 Db 62 KCVGGRIKH 71

RESULT 12

O8T475 PRELIMINARY; PRT; 119 AA.

AC O8T475; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

GN AT12574P.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacleob J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celiker S.;

RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY089320; AAL90058.1; -;

FT SEQUENCE 119 AA; 13300 MW; 3FF85BC91F8068 CRC64;

Query Match 50.0%; Score 46; DB 5; Length 119;

Best Local Similarity 80.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 || || || || ||
 DB 66 ECPGGGRIEH 75

RESULT 13

O9BM98 PRELIMINARY; PRT; 135 AA.
 ID O9BM98
 AC O9BM98
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janusa.
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Melkilejohn C.D., Hauschteck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
 RT melanogaster Species Subgroup.";
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL; AY013340; AAG50361.1; -;
 DR Flybase; FBgn0043691; Dmau\Jana.
 SQ SEQUENCE 135 AA; 15158 MW; 2F91E387BD70C1CC CRC64;

Query Match

Best Local Similarity 50.0%; Score 46; DB 5; Length 135;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 || || || || ||
 DB 82 ECPGGGRIEH 91

RESULT 14

O9BM97 PRELIMINARY; PRT; 135 AA.
 ID O9BM97
 AC O9BM97;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janusa.
 GN JANA.
 OS Drosophila erecta (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Melkilejohn C.D., Hauschteck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
 RT melanogaster Species Subgroup.";
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL; AY013343; AAG50364.1; -;
 DR Flybase; FBgn0043709; Dore\Jana.
 SQ SEQUENCE 135 AA; 15180 MW; DB294031FEDFD07C CRC64;

Query Match

Best Local Similarity 50.0%; Score 46; DB 5; Length 135;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 || || || || ||
 DB 82 ECPGGGRIEH 91

RESULT 15

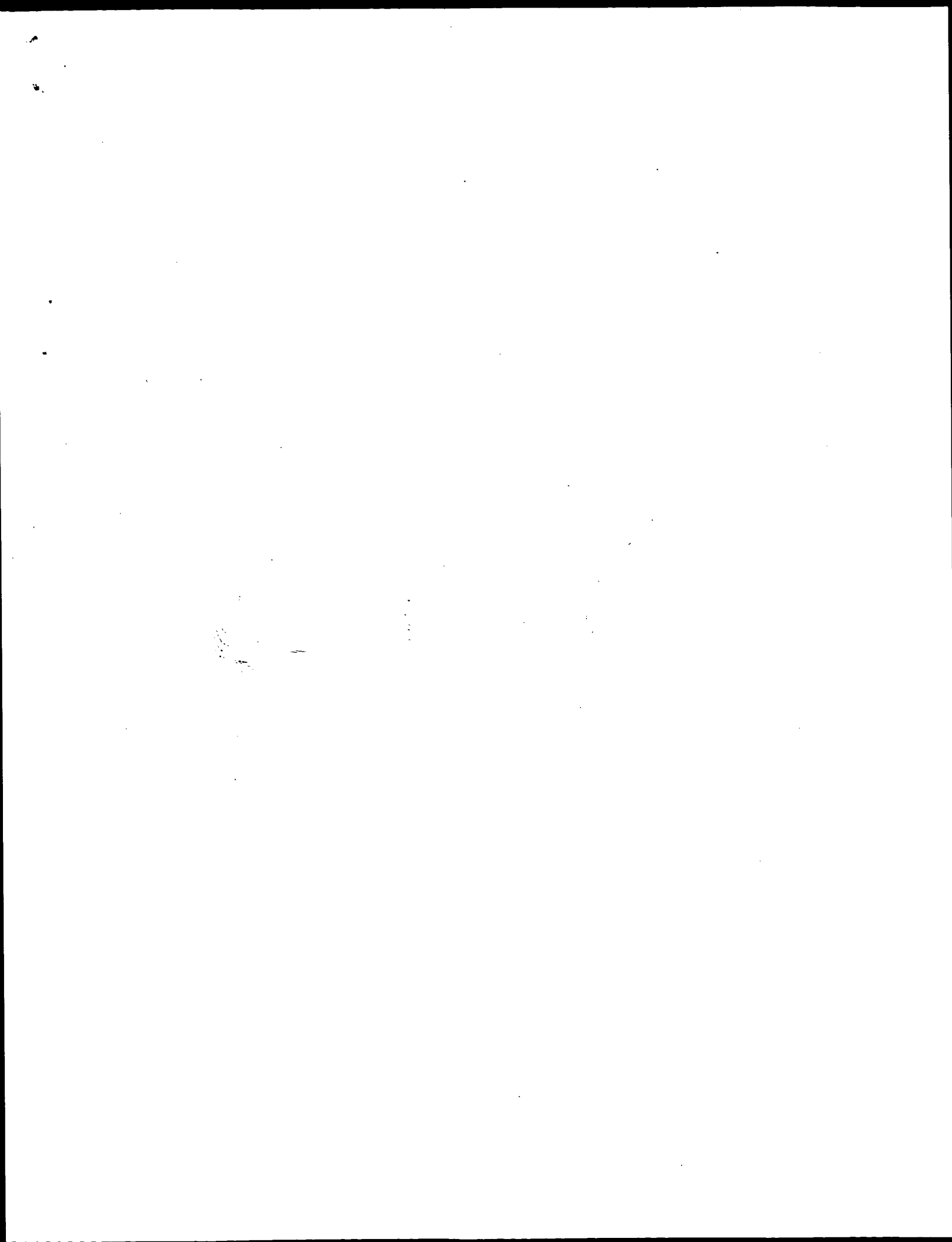
O9BM96 PRELIMINARY; PRT; 135 AA.
 ID O9BM96
 AC O9BM96;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Janusa.
 GN JANA.
 OS Drosophila oreana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Melkilejohn C.D., Hauschteck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
 RT melanogaster Species Subgroup.";
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL; AY013344; AAG50365.1; -;
 DR Flybase; FBgn0043667; Dore\Jana.
 SQ SEQUENCE 135 AA; 15146 MW; DB21E237BCD1307C CRC64;

Query Match

Best Local Similarity 50.0%; Score 46; DB 5; Length 135;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ECLGGGRISH 12
 || || || || ||
 DB 82 ECPGGGRIEH 91

Search completed: May 7, 2003, 19:06:38
 Job time : 14.0826 secs



GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:25 ; Search time 4.25688 Seconds

(without alignments)
110.590 Million cell updates/sec

Title: US-09-914-831-3

Perfect score: 92

Sequence: 1 DCECLGGGRISHOSQD 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	50.0	48	4	US-08-438-745-22
2	46	50.0	48	5	PCT-US94-05669A-22
3	44	47.8	353	4	US-09-171-156A-65
4	44	47.8	353	4	US-09-171-156A-65
5	44	47.8	548	4	US-09-167-299-3
6	42	45.7	1229	1	US-08-100-709-4
7	42	45.7	1229	1	US-08-176-865-4
8	42	45.7	1229	1	US-08-474-038-4
9	42	45.7	1229	2	US-08-779-046-4
10	42	45.7	1229	2	US-08-881-340-4
11	42	45.7	2231	1	US-08-153-799-16
12	42	45.7	2324	1	US-08-283-857-1
13	42	45.7	2324	1	US-08-283-857-1
14	42	45.7	2327	6	PCT-US95-09819-1
15	42	45.7	2386	2	US-09-016-366A-12
16	42	45.7	2446	2	US-08-551-356-2
17	42	45.7	2446	5	PCT-US93-12687-2
18	40	43.5	149	4	US-09-183-861-86
19	40	43.5	149	4	US-09-183-861-86
20	39.5	42.9	115	2	US-08-485-449-4
21	39.5	42.9	376	2	US-08-485-449-5
22	39.5	42.9	389	2	US-08-485-449-6
23	39.5	42.9	389	2	US-08-485-449-7
24	39.5	42.9	389	2	US-08-485-449-7
25	39	42.4	432	2	US-08-731-079A-2
26	39	42.4	934	1	US-08-446-486-7
27	39	42.4	934	1	US-08-463-308-7

28	39	42.4	969	1	US-07-671-817A-4	Sequence 4, Appl
29	39	42.4	969	1	US-07-671-817A-6	Sequence 6, Appl
30	39	42.4	1129	6	5164180-6	Patent No. 5164180
31	39	42.4	1167	1	US-08-100-709-2	Sequence 2, Appl
32	39	42.4	1167	1	US-08-176-865-2	Sequence 2, Appl
33	39	42.4	1167	1	US-08-474-038-2	Sequence 2, Appl
34	39	42.4	1167	2	US-08-779-046-2	Sequence 2, Appl
35	39	42.4	1167	2	US-08-881-340-2	Sequence 2, Appl
36	39	42.4	1168	1	US-08-448-170-6	Sequence 6, Appl
37	39	42.4	1168	4	US-08-961-803-8	Sequence 8, Appl
38	39	42.4	1170	1	US-08-032-364-2	Sequence 2, Appl
39	39	42.4	1171	1	US-07-828-788A-6	Sequence 6, Appl
40	39	42.4	1171	1	US-08-278-685-2	Sequence 2, Appl
41	39	42.4	1171	1	US-08-277-721-2	Sequence 2, Appl
42	39	42.4	1171	1	US-08-277-721-2	Sequence 4, Appl
43	39	42.4	1171	1	US-08-602-737-4	Sequence 4, Appl
44	39	42.4	1171	1	US-09-001-982-4	Sequence 4, Appl
45	39	42.4	1171	5	PCT-US92-11337-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-438-745-22
Sequence 22, Application US/08438745
Patent No. 6248715
GENERAL INFORMATION:
APPLICANT: Rosenberg, Steven
APPLICANT: Stratton-Thomas, Jennifer
TITLE OF INVENTION: Expression of urokinase Plasminogen
TITLE OF INVENTION: Activator Inhibitors
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,745
FILING DATE: 10-MAY-1995
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/070,153
FILING DATE: 01-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939, 001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-438-745-22

Query Match 50.0%; Score 46; DB 4; Length 48;
Best Local Similarity 53.8%; Pred. No. 1.9;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHQ 13
:|:|:|:|:|:|
DB 10 NCDCLNGTGVHQ 22

RESULT 2

PCT-US94-05669A-22
Sequence 22, Application PC/TUS9405669A
GENERAL INFORMATION:
APPLICANT: Chiron Corporation
TITLE OF INVENTION: Expression of Urokinase Plasminogen
Activator Inhibitors
TITLE OF INVENTION: Activator Inhibitors
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05669A
FILING DATE: 19-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D. 31,259
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939,100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05669A-22

Query Match 50.0%; Score 46; DB 5; Length 48;
Best Local Similarity 53.8%; Pred. No. 1.9;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DCECLGGGRISHQ 13
:|:|:|:|:|:|
DB 10 NCDCLNGTGVHQ 22

RESULT 3

US-09-171-156A-65
Sequence 65, Application US/09171156A
Patent No. 6368846
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
Slm, Gek-kee
Weber, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO

COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-09-171-156A-65

Query Match 47.8%; Score 44; DB 4; Length 353;
Best Local Similarity 46.7%; Pred. No. 28;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCECLGGGRISHOSQD 16
|:|:|:|:|:|:|
DB 17 CDYGPGRKITHKESD 31

RESULT 4

US-09-171-156A-68
Sequence 68, Application US/09171156A
Patent No. 6368846
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
Slm, Gek-kee
Weber, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids

Best Local Similarity 50.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 6;

OY 1 DCECLGGGRISHOS 14
DB 863 DCSRDGERCAHNS 876

RESULT 8
US-08-474-038-4
; Sequence 4, Application US/08474038
; Patent No. 5679343
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panlitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,038
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,865
; FILING DATE: 30-DEC-1993
; APPLICATION NUMBER: US 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-474-038-4
Query Match 45.7%; Score 42; DB 1; Length 1229;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
OY 1 DCECLGGGRISHOS 14
DB 863 DCSRDGERCAHNS 876
RESULT 9
US-08-779-046-4
; Sequence 4, Application US/08779046
; Patent No. 5854053
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.

APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panlitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,046
; FILING DATE: 06-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-779-046-4
Query Match 45.7%; Score 42; DB 2; Length 1229;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
OY 1 DCECLGGGRISHOS 14
DB 863 DCSRDGERCAHNS 876
RESULT 10
US-08-881-340-4
; Sequence 4, Application US/08881340
; Patent No. 5942658
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panlitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,340
FILING DATE: 24-JUN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,709
FILING DATE: 29-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Egoft, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-49
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1229 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-861-340-4

Query Match 45.7%; Score 42; DB 2; Length 1229;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 DCECIGGGRSHOS 14
Db 863 DSCRCGRCAHNS 876

RESULT 11
US-08-153-799-16
Sequence 16, Application US/08153799
Patent No. 5766883
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 928832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159

TELEX: 219484
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /note="Human fibronectin"
US-08-153-799-16

Query Match 45.7%; Score 42; DB 1; Length 2231;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCECIGGGR 9
Db 91 DCTCIGAGR 99

RESULT 12
US-08-283-857-1
Sequence 1, Application US/08283857
Patent No. 5792742
GENERAL INFORMATION:
APPLICANT: GOLD, Leslie I.
APPLICANT: ROSTAGNO, Agueda A.
APPLICANT: BARON, Martin
APPLICANT: CAMPBELL, Iain D.
APPLICANT: WILLIAMS, Michael, J.
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,857
FILING DATE: 01-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/714,134
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-283-857-1

Query Match 45.7%; Score 42; DB 1; Length 2324;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCECIGGR 9
Db 91 DCTCIGGR 99

RESULT 13

PCT-US95-09819-1
Sequence 1, Application PC/TUS9509819
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09819
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,857
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match 45.7%; Score 42; DB 5; Length 2324;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCECIGGR 9
Db 91 DCTCIGGR 99

RESULT 14
5455158-1
Patent No. 5455158
APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WEBER, MOSHE M.;
GUY, RACHEL; PANET, AMOS
TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
USES AND METHODS OF PRODUCING SAME
NUMBER OF SEQUENCES: 20
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/58,241
FILING DATE: 04-MAY-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 526,397
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: 345,952
FILING DATE: 28-APR-1989
APPLICATION NUMBER: 291,951
FILING DATE: 29-DEC-1988
SEQ ID NO: 1
LENGTH: 2327
5455158-1

Query Match 45.7%; Score 42; DB 6; Length 2327;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DCECIGGR 9
Db 94 DCTCIGGR 102

RESULT 15

US-09-016-366A-12
Sequence 12, Application US/09016366A
Patent No. 595431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2386 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-016-366A-12

Query Match 45.7%; Score 42; DB 2; Length 2386;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

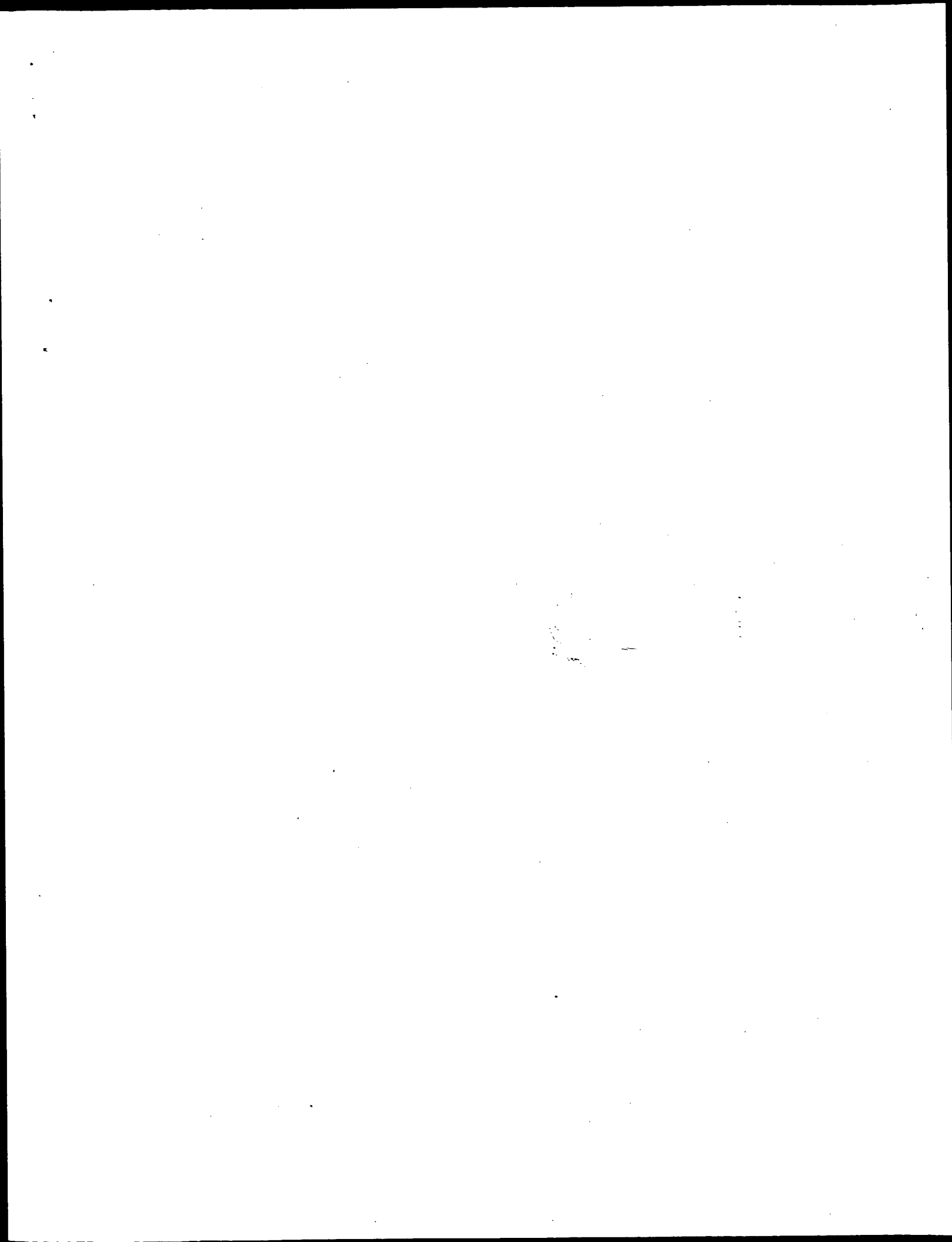
OY 1 DCECIGGR 9
Db 122 DCTCIGGR 130

Wed May 14 09:22:08 2003.

us-09-914-831-3.rai

Page 7

Search completed: May 7, 2003, 19:08:36
Job time : 6.25688 secs



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 18:54:20 ; Search time 23.4128 Seconds
(Without alignments)
250,419 Million cell updates/sec

Title: US-09-914-831-5

Perfect score: 249
Sequence: 1 YHADIVKVSQDMQKQDC.....GGRISHQSDKKIHVGYSM 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249	100.0	44	21	AA08790
2	249	100.0	125	21	AA08787
3	249	100.0	125	22	AA03661
4	249	100.0	165	22	AA04834
5	249	100.0	165	21	AA04276
6	249	100.0	165	22	AA04447
7	223	89.6	124	21	AA08791
8	220	88.4	124	21	AA08793
9	208	83.5	123	21	AA08792
10	141	56.6	33	21	AA08789

11	109	43.8	135	22	AB058764
12	101	40.6	119	22	AA031440
13	92	36.9	16	21	AA08788
14	92	36.9	140	22	AB058770
15	90	36.1	160	22	AB067163
16	88	35.3	148	22	AB058760
17	77	30.9	14	21	AA08795
18	60.5	24.3	918	22	AB05892
19	58	23.3	1443	23	AB093701
20	55.5	22.3	130	22	AA096394
21	55	22.1	136	22	AA003501
22	55	22.1	444	21	AA027888
23	54.5	21.9	259	23	AB054545
24	54	21.7	149	19	AA070257
25	54	21.7	149	23	AB069090
26	54	21.7	149	23	AA071853
27	54	21.7	611	22	AB019321
28	53.5	21.5	249	20	AA086310
29	53	21.3	130	22	AA092194
30	53	21.3	284	21	AA031468
31	53	21.3	319	21	AA031467
32	53	21.3	345	21	AA031466
33	53	21.3	351	22	AB027382
34	53	21.3	351	22	AB032530
35	53	21.3	351	22	AB018034
36	53	21.3	351	22	AA053363
37	53	21.3	351	22	AA065742
38	53	21.3	351	22	AA013601
39	53	21.3	351	22	AA025998
40	53	21.3	351	22	AA001351
41	53	21.3	352	21	AB035371
42	53	21.3	352	21	AA080516
43	53	21.3	372	21	AA017037
44	53	21.3	372	21	AA042989
45	53	21.3	373	21	AA035271

ALIGNMENTS

RESULT 1
ID AAB08790 standard; Peptide: 44 AA.
XX AAB08790;
AC
XX
DT 02-JAN-2001 (first entry)
XX
DE Conserved sequence of mammalian histidine protein phosphatases.
XX
XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Mammalia sp.
XX
XX WO200052175-A1.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-EP01774.
XX
XX 04-MAR-1999; 99DE-1009388.
XX
XX (MERE) MERCK PATENT GMBH.
XX
XX Klump S, Kellner R;
XX
XX WPI: 2000-572187/53.
XX
XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
XX immune disorders, viral infection, genetic disorders, and heart disease
XX

XX Claim 4; Page 17; 39pp; English.

PS The present sequence represents a conserved sequence found in mammalian

CC histidine phosphatase polypeptides. The human polypeptide has a high

CC specificity for phosphohistidine and a molecular weight of 13000-15000

CC Da. The human histidine phosphatase gene is localised at chromosome 9

CC (9q33). The human histidine phosphatase can be used for diagnosis and

CC treatment of pathological states of cell regulation and cell growth.

CC These include cancers, immune disorders, viral infection, genetic

CC disorders, and heart disease. The histidine phosphatase can also be

CC used for identifying agonists and antagonists which can be used to

CC treat conditions associated with N-phosphorylation imbalance.

XX

SQ Sequence 44 AA;

Query Match 100.0%; Score 249; DB 21; Length 44;

Best Local Similarity 100.0%; Pred. No. 1.8e-24;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YHADYKVSQDMQKQCCDECGGRISHOSQDKRIHYGYSM 44

DB 1 YHADYKVSQDMQKQCCDECGGRISHOSQDKRIHYGYSM 44

RESULT 2

ID AAB08787 standard; Protein; 125 AA.

XX AAB08787;

XX 02-JAN-2001 (first entry)

DE A human histidine protein phosphatase polypeptide.

XX

XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;

KW cell regulation; cell growth; cancer; immune disorder; viral infection;

KW genetic disorder; heart disease; N-phosphorylation imbalance.

XX

XX Homo sapiens.

XX WO200052175-A1.

XX

XX 08-SEP-2000.

XX

XX 02-MAR-2000; 2000WO-EP01774.

XX

XX 04-MAR-1999; 99DE-1009388.

XX

XX (MERE) MERCK PATENT GMBH.

XX

XX Klump S, Kellner R;

XX WPI; 2000-572187/53.

DR N-PSDB; AAA64502.

XX

PT Histidine phosphatase, useful for diagnosis and treatment of cancers,

PT immune disorders, viral infection, genetic disorders, and heart disease

XX

XX Claim 5; Page 17; 39pp; English.

XX

XX The present sequence represents human histidine phosphatase polypeptide.

CC The polypeptide has a high specificity for phosphohistidine and a

CC molecular weight of 13000-15000 Da. The histidine phosphatase gene is

CC localised at chromosome 9 (9q33). The histidine phosphatase can be used

CC for diagnosis and treatment of pathological states of cell regulation

CC and cell growth. These include cancers, immune disorders, viral

CC infection, genetic disorders, and heart disease. The histidine

CC phosphatase can also be used for identifying agonists and antagonists

CC which can be used to treat conditions associated with N-phosphorylation

CC imbalance.

XX

SQ Sequence 125 AA;

Query Match 100.0%; Score 249; DB 21; Length 125;

Best Local Similarity 100.0%; Pred. No. 5.7e-24;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YHADYKVSQDMQKQCCDECGGRISHOSQDKRIHYGYSM 44

DB 52 YHADYKVSQDMQKQCCDECGGRISHOSQDKRIHYGYSM 95

RESULT 3

ID AAM39661 standard; Protein; 125 AA.

XX AAM39661;

XX 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2806.

XX

XX Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX

XX Homo sapiens.

XX WO200153312-A1.

XX

XX 26-JUL-2001.

XX

XX 26-DEC-2000; 2000WO-US34263.

XX

XX 21-JAN-2000; 2000US-0488725.

XX

XX 25-APR-2000; 2000US-0552317.

XX

XX 09-JUL-2000; 2000US-0598042.

XX

XX 19-JUL-2000; 2000US-0620312.

XX

XX 03-AUG-2000; 2000US-0653450.

XX

XX 14-SEP-2000; 2000US-0662191.

XX

XX 19-OCT-2000; 2000US-0693036.

XX

XX 29-NOV-2000; 2000US-0727344.

XX

XX (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

XX WPI; 2001-442253/47.

DR N-PSDB; AAI58817.

XX

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

XX Example 4; SEQ ID NO 2806; 10078pp; English.

XX

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAI42213) with nocotropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 125 AA;

Query Match 100.0%; Score 249; DB 22; Length 125;
 Best Local Similarity 100.0%; Pred. No. 5,7e-24;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YHADIYDKVSGDMKQKCGDCCECLGGGRISHOSQDKIHVGYSM 44
 |||
 DB 52 YHADIYDKVSGDMKQKCGDCCECLGGGRISHOSQDKIHVGYSM 95

RESULT 4

AA48348
 ID AA48348 standard; Protein: 165 AA.

AC AA48348;

DT 08-DEC-1999 (first entry)

DE Human prostate cancer-associated protein 45.

KW Expressed sequence tag: EST; prostate; tumor; treatment; gene therapy;
 KW cancer; tissue specificity; human.

OS Homo sapiens.

FN DEJ9811194-A1.

PD 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011194.

PR 10-MAR-1998; 98DE-1011194.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI: 1999-519629/44.

DR N-PSDB; AA233497.

PT New nucleic acid expressed at high level in normal prostatic tissue and
 PT encoded polypeptides, used to treat cancer and screen for therapeutic
 PT agents

PS Claim 22: 141; 194pp; German.

XX This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in normal prostatic tissue. Polypeptides (1)
 CC encoded by (A) are used: (a) for identifying agents for treatment of
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally
 CC where expressed by gene therapy methods. (A) is also used to isolate
 CC full-length genes (for gene therapy) and for recombinant production of
 CC (1), which can be used to raise specific antibodies. (A) are identified
 CC by assembly of ESTs (expressed sequence tags) before these are analyzed
 CC many of the false results, as regards tissue specificity. This approach eliminates
 CC many of the false results, as regards tissue specificity, associated
 CC with known methods that use single (usually short) ESTs. AA48304-48456
 CC represent peptides encoded by the expressed sequence tags described in
 CC the method of the invention.

XX Sequence 165 AA;

Query Match 100.0%; Score 249; DB 20; Length 165;
 Best Local Similarity 100.0%; Pred. No. 7.8e-24;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YHADIYDKVSGDMKQKCGDCCECLGGGRISHOSQDKIHVGYSM 44
 |||
 DB 92 YHADIYDKVSGDMKQKCGDCCECLGGGRISHOSQDKIHVGYSM 135

RESULT 5

AA42776
 ID AA42776 standard; Protein: 165 AA.

AC AA42776;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2540 polypeptide sequence SEQ ID NO:5080.

KW Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerability; antiproliferative; antiparkinsonian; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; aschma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

OS Homo sapiens.

FN WO200058473-A2.

PD 05-OCT-2000.

PF 31-MAR-2000; 2000MO-US08621.

PR 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

PA (CURA-) CURAGEN CORP.

PI Shimkets RA, Leach M;

DR WPI: 2000-602362/57.

DR N-PSDB; AAC76985.

PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

PS Claim 11; Page 4264-4265; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40337 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antirheumatic; antihypoid;
 CC antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 CC neurodegenerative disorder; osteoarthritis; graft vs host disease;
 CC cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 CC cholesterol ester storage; systemic lupus erythematosus; infection;
 CC severe combined immunodeficiency (SCID); AIDS; viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance

CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 CC Sequence 165 AA;

Query Match 100.0%; Score 249; DB 21; Length 165;
 Best Local Similarity 100.0%; Pred. No. 7.8e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YHADIYKVSQDMKQGCDCGCGRISHQSDKKIHYGYSM 44
 ||||||||||||||||||||||||||||||||||||||||
 Db 92 YHADIYKVSQDMKQGCDCGCGRISHQSDKKIHYGYSM 135

RESULT 6

AAM41447
 ID AAM41447 standard; Protein; 165 AA.

AC AAM41447;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6378.

XX Human; nocitropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0532317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-ANG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.

DR N-PSDB: AAI60603.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6378; 10078bp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nocitropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 165 AA;

Query Match 100.0%; Score 249; DB 22; Length 165;
 Best Local Similarity 100.0%; Pred. No. 7.8e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YHADIYKVSQDMKQGCDCGCGRISHQSDKKIHYGYSM 44
 ||||||||||||||||||||||||||||||||||||||||
 Db 92 YHADIYKVSQDMKQGCDCGCGRISHQSDKKIHYGYSM 135

RESULT 7

AAB08791
 ID AAB08791 standard; Peptide; 124 AA.

AC AAB08791;

DT 02-JAN-2001 (first entry)

DE Rabbit histidine protein phosphatase polypeptide.

XX Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
 KW cell regulation; cell growth; cancer; immune disorder; viral infection;
 KW genetic disorder; heart disease; N-phosphorylation imbalance.

OS Oryctolagus cuniculus.

PN WO200052175-A1.

PD 08-SEP-2000.

PF 02-MAR-2000; 2000WO-EP01774.

PR 04-MAR-1999; 99DE-1009388.

PA (MERE) MERCK PATENT GMBH.

PI Klump S, Kellner R;

DR WPI: 2000-572187/53.

XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
 PT immune disorders, viral infection, genetic disorders, and heart disease

PS Claim 7; Page 18; 39pp; English.

XX The present sequence represents a rabbit histidine phosphatase
 CC polypeptide. The human polypeptide has a high specificity for
 CC phosphohistidine and a molecular weight of 13000-15000 Da. The human
 CC histidine phosphatase gene is localised at chromosome 9 (9q33). The
 CC human histidine phosphatase can be used for diagnosis and treatment
 CC of pathological states of cell regulation and cell growth. These
 CC include cancers, immune disorders, viral infection, genetic disorders,
 CC and heart disease. The histidine phosphatase can also be used for
 CC identifying agonists and antagonists which can be used to treat
 CC conditions associated with N-phosphorylation imbalance.

XX Sequence 124 AA;

Query Match 89.6%; Score 223; DB 21; Length 124;
 Best Local Similarity 88.6%; Pred. No. 1.1e-20;
 Matches 39; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 YHADIYKVSQDMKQGCDCGCGRISHQSDKKIHYGYSM 44
 ||||||||||||||||||||||||||||||||||||||||
 Db 51 YHADIYKVSQDMKQGCDCGCGRISHQSDKKIHYGYSM 94

RESULT 8

AAB08793

ID AAB08793 standard; Peptide: 124 AA.

XX AAB08793;

DT 02-JAN-2001 (first entry)

DE Mouse histidine protein phosphatase polypeptide.

XX Human: histidine phosphatase; phosphohistidine; chromosome 9; 9q33;

XX cell regulation; cell growth; cancer; immune disorder; viral infection;

XX genetic disorder; heart disease; N-phosphorylation imbalance.

XX OS Mus sp.

XX WO20052175-A1.

XX 08-SEP-2000.

XX 02-MAR-2000; 2000WO-EP01774.

XX 04-MAR-1999; 99DE-1009388.

XX (MERE) MERCK PATENT GMBH.

XX Klump S, Kellner R;

XX WPI: 2000-572187/53.

XX Histidine phosphatase, useful for diagnosis and treatment of cancers,

XX immune disorders, viral infection, genetic disorders, and heart disease

XX -

XX

XX Disclosure; Page 12; 39pp; English.

XX The present sequence represents a mouse histidine phosphatase

XX polypeptide. The human polypeptide has a high specificity for

XX phosphohistidine and a molecular weight of 13000-15000 Da. The human

XX histidine phosphatase gene is localised at chromosome 9 (9q33). The

XX human histidine phosphatase can be used for diagnosis and treatment

XX of pathological states of cell regulation and cell growth. These

XX include cancers, immune disorders, viral infection, genetic disorders,

XX and heart disease. The histidine phosphatase can also be used for

XX identifying agonists and antagonists which can be used to treat

XX conditions associated with N-phosphorylation imbalance.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

OS Rattus sp.

XX WO20052175-A1.

XX 08-SEP-2000.

DT 02-MAR-2000; 2000WO-EP01774.

DE 04-MAR-1999; 99DE-1009388.

XX (MERE) MERCK PATENT GMBH.

XX Klump S, Kellner R;

XX WPI: 2000-572187/53.

XX Histidine phosphatase, useful for diagnosis and treatment of cancers,

XX immune disorders, viral infection, genetic disorders, and heart disease

XX -

XX Disclosure; Page 12; 39pp; English.

XX The present sequence represents a rat histidine phosphatase

XX polypeptide. The human polypeptide has a high specificity for

XX phosphohistidine and a molecular weight of 13000-15000 Da. The human

XX histidine phosphatase gene is localised at chromosome 9 (9q33). The

XX human histidine phosphatase can be used for diagnosis and treatment

XX of pathological states of cell regulation and cell growth. These

XX include cancers, immune disorders, viral infection, genetic disorders,

XX and heart disease. The histidine phosphatase can also be used for

XX identifying agonists and antagonists which can be used to treat

XX conditions associated with N-phosphorylation imbalance.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX 04-MAR-1999; 99DE-1009388.
 XX (MERE) MERCK PATENT GMBH.
 XX
 XX Klumpp S, Kellner R;
 XX WPI: 2000-572187/53.
 XX
 XX Histidine phosphatase, useful for diagnosis and treatment of cancers,
 XX immune disorders, viral infection, genetic disorders, and heart disease
 XX
 XX Claim 3; Page 17; 39pp; English.
 XX
 XX The present sequence represents a conserved sequence found in mammalian
 XX histidine phosphatase polypeptides. The human polypeptide has a high
 XX specificity for phosphohistidine and a molecular weight of 13000-15000
 XX Da. The human histidine phosphatase gene is localized at chromosome 9
 XX (9q33). The human histidine phosphatase can be used for diagnosis and
 XX treatment of pathological states of cell regulation and cell growth.
 XX These include cancers, immune disorders, viral infection, genetic
 XX disorders, and heart disease. The histidine phosphatase can also be
 XX used for identifying agonists and antagonists which can be used to
 XX treat conditions associated with N-phosphorylation imbalance.
 XX
 XX Sequence 33 AA;
 XX
 XX Query Match 56.6%; Score 141; DB 21; Length 33;
 XX Best Local Similarity 96.2%; Pred. No. 6.8e-11;
 XX Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX 19 DCECLGGGRISHOSQDKRIHYGYSM 44
 XX |||||
 XX 1 DCECLGGGRISHOSQDKRIHYGYSM 26
 XX
 XX RESULT 11
 XX ABB58764
 XX ID ABB58764 standard; Protein: 135 AA.
 XX
 XX ABB58764;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 3084.
 XX
 XX Drosophila: developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX
 XX Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 XX
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PMD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 XX
 XX N-PSDB; ABLO2867.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -
 XX
 XX Disclosure; SEQ ID NO 3084; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 XX sequences (AB101840-AB16175) and the encoded proteins
 XX (AB57737-AB572072).
 XX The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 135 AA;
 XX
 XX Query Match 43.8%; Score 109; DB 22; Length 135;
 XX Best Local Similarity 46.5%; Pred. No. 3.7e-06;
 XX Matches 20; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
 XX
 XX 1 YHADIYKVSQDMQKQDCDECLGGGRISHOSQDKRIHYGYS 43
 XX :|||||: :| | | | | : : : | | | |
 XX 62 WHADIFREEVKTKGLRAECPCGGRIEHNPCKYLYKYGYS 104
 XX
 XX RESULT 12
 XX AAU31440
 XX ID AAU31440 standard; Protein: 119 AA.
 XX
 XX AAU31440;
 XX
 XX 18-DEC-2001 (first entry)
 XX
 XX Novel human secreted protein #1931.
 XX
 XX Human: vaccination; gene therapy; nutritional supplement;
 XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 XX Homo sapiens.
 XX
 XX WO200179449-A2.
 XX
 XX 25-OCT-2001.
 XX
 XX 16-APR-2001; 2001WO-US08656.
 XX
 XX 18-APR-2000; 2000US-0552929.
 XX
 XX 26-JAN-2001; 2001US-0770160.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 XX
 XX WPI: 2001-611725/70.
 XX
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy -
 XX
 XX Claim 20; Page 440; 765pp; English.
 XX
 XX The invention relates to novel human secreted polypeptides. The
 XX polypeptides and antibodies to the polypeptides are useful for
 XX determining the presence of or predisposition to a disease associated
 XX with altered levels of polypeptide. The polypeptides are also useful for
 XX identifying agents (agonists and antagonists) that bind to them. Cells
 XX expressing the proteins are useful for identifying a therapeutic agent
 XX for use in treatment of a pathology related to aberrant expression or
 XX physiological interactions of the polypeptide. Vectors comprising
 XX the nucleic acids encoding the polypeptides and cells genetically
 XX engineered to express them are also useful for producing the proteins.
 XX The proteins are useful in genetic vaccination, testing and
 XX therapy, and can be used as nutritional supplements. They may be used to
 XX increase stem cell proliferation; to regulate haematopoiesis; and in

CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 119 AA;

Query Match 40.6%; Score 101; DB 22; Length 119;
Best Local Similarity 52.1%; Pred. No. 3.4e-05;
Matches 25; Conservative 4; Mismatches 15; Indels 4; Gaps 3;

OY 1 YHADIYDK-VSGDMQKGC-DCEELGGGRISHOSQ--DKRIHYGYSM 44
Db 60 HHADIDKSVGATCRKQCLRTVSLGGGRISHSRKSPQGRFTYGYSM 107

RESULT 13
AAB08788
ID AAB08788 standard; Peptide; 16 AA.
XX
AC AAB08788;

02-JAN-2001 (first entry)

Conserved sequence of mammalian histidine protein phosphatases.

KW Human; histidine phosphatase; phosphohistidine; chromosome 9; 9q33;
KW cell regulation; cell growth; cancer; immune disorder; viral infection;
KW genetic disorder; heart disease; N-phosphorylation imbalance.
XX
OS Mammalia sp.

PN MO200052175-A1.

PD 08-SEP-2000.

PF 02-MAR-2000; 2000MO-EP01774.

PR 04-MAR-1999; 99DE-1009388.

PA (MERCK) MERCK PATENT GMBH.

PI Klumpner S, Kellner R;

DR WPI: 2000-572187/53.

PT Histidine phosphatase, useful for diagnosis and treatment of cancers,
PT immune disorders, viral infection, genetic disorders, and heart disease
XX
PS Claim 2; Page 17; 39pp; English.

CC The present sequence represents a conserved sequence found in mammalian
CC histidine phosphatase polypeptides. The human polypeptide has a high
CC specificity for phosphohistidine and a molecular weight of 13000-15000
CC Da. The human histidine phosphatase gene is localised at chromosome 9
CC (9q33). The human histidine phosphatase can be used for diagnosis and
CC treatment of pathological states of cell regulation and cell growth.
CC These include cancers, immune disorders, viral infection, genetic
CC disorders, and heart disease. The histidine phosphatase can also be
CC used for identifying agonists and antagonists which can be used to
CC treat conditions associated with N-phosphorylation imbalance.
XX
SQ Sequence 16 AA;

Query Match 36.9%; Score 92; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 DCECGGGRISHOSQD 34
Db 1 DCECLGGGRISHOSQD 16

RESULT 14
ABB58770
ID ABB58770 standard; Protein; 140 AA.
XX
AC ABB58770;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 3102.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.

PN MO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-C614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI: 2001-656860/75.

DR N-PSDB; ABL02873.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 3102; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 140 AA;

Query Match 36.9%; Score 92; DB 22; Length 140;
Best Local Similarity 35.7%; Pred. No. 0.00056;
Matches 15; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

OY 2 HADIDKSVGDMQKGC-DCEELGGGRISHOSQDKRIHYGYSM 43
Db 69 HLAVFDLLELEPBGICAKILGGGRILNENMKRIKITYGS 110

RESULT 15

ABB67163
ID ABB67163 standard; Protein; 160 AA.

AC ABB67163;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 28281.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

```

XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.
XX DR N-PSDB; ABL11266.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 28281; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 160 AA;
Query Match 36.1%; Score 90; DB 22; Length 160;
Best Local Similarity 37.2%; Pred. No. 0.0012;
Matches 16; Conservative 9; Mismatches 18; Indels 0; Gaps 0;
OY 1 YHADIDYKVSQDMQKQDCGCGRISHSQSDKKIHVYGS 43
| | | | | : : : | | | | : : | | |
DB 85 YHIDVYDELKKEAEMSGICTQGLGGYIVHDKKRYIKLYGRS 127

```

Search completed: May 7, 2003, 19:03:27
 Job time : 24.4128 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:01:00 ; Search time 9.48624 Seconds

(without alignments)
445,900 Million cell updates/sec

Title: US-09-914-831-5

Sequence: 1 YHADIYDKVSGDMKQKGCDC.....GGRISHQSDKKIHVGYGS 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	46.2	149	2	S53876 sex-regulated prot
2	109	43.8	135	2	A32317 sex-regulated prot
3	92	36.9	140	2	B32317 sex-regulated prot
4	78.5	31.5	115	2	T21833 hypothetical prote
5	77	30.9	140	2	S53877 sex-regulated prot
6	60.5	24.3	548	2	S20806 swallow protein -
7	60	24.1	477	2	T33531 hypothetical prote
8	60	24.1	906	2	D82899 alanyl-tRNA synth
9	58.5	23.5	382	2	F81298 probable decarboxy
10	57	22.9	444	2	S31191 tubulin beta-1 cha
11	57	22.9	1594	2	T43072 hemolysin A - Edwa
12	56.5	22.7	296	1	WMAD15 late 100k protein
13	55.5	22.7	807	1	MMAD15 late 100k protein
14	55.5	22.3	430	2	A65215 hypothetical 49.4
15	55	22.1	204	2	S02532 tubulin beta-1 cha
16	55	22.1	282	2	A25342 tubulin beta chain
17	55	22.1	444	2	A35885 tubulin beta chain
18	55	22.1	445	2	A44848 beta 1A tubulin -
19	55	22.1	451	2	S30514 tubulin beta chain
20	54.5	21.9	244	2	B47754 cyclase hsf - Lac
21	54.5	21.9	259	2	F86776 cyclase hsf [limp
22	54	21.7	442	2	S25182 tubulin beta-1 cha
23	54	21.7	442	2	S00683 tubulin beta-1 cha
24	54	21.7	443	2	S01769 tubulin beta-1 cha
25	54	21.7	443	2	S01768 tubulin beta-1 cha
26	54	21.7	443	2	S41470 tubulin beta-2 cha
27	54	21.7	443	2	S32668 tubulin beta chain
28	54	21.7	444	2	S31400 tubulin beta chain
29	54	21.7	445	2	A49494 tubulin beta chain

30	54	21.7	445	2	JA0048 tubulin beta-1 cha
31	54	21.7	445	2	A54515 tubulin beta chain
32	54	21.7	446	2	B30309 tubulin beta chain
33	54	21.7	449	2	S16340 tubulin beta chain
34	53.5	21.5	162	2	B29662 Balbiani ring 2 ch
35	53.5	21.5	359	2	B29960 Balbiani ring 2 ch
36	53.5	21.5	359	2	B29960 Balbiani ring 2 ch
37	53.5	21.5	1503	2	S17983 gene posterior sex
38	53	21.3	345	2	C96632 hypothetical prote
39	53	21.3	383	2	S42480 tubulin beta chain
40	53	21.3	386	2	S14570 tubulin beta chain
41	53	21.3	411	2	S32669 tubulin beta-2 cha
42	53	21.3	412	2	T40155 manose-6-phosphat
43	53	21.3	434	2	S52008 tubulin beta-2 cha
44	53	21.3	437	2	I38369 tubulin beta-3 cha
45	53	21.3	440	2	S20870 tubulin beta-3 cha

ALIGNMENTS

RESULT 1

S53876 sex-regulated protein janus A - fruit fly (Drosophila pseudoobscura)

C:Species: Drosophila pseudoobscura

C:Date: 14-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Jan-1999

C:Accession: S53876

R:Yanicosas, C.; Ferrer, P.; Vincent, A.; Lepeant, J.A.

Mol. Gen. Genet. 246, 549-560, 1995

A:Title: Separate cis-regulatory sequences control expression of serendipity beta and

A:Reference number: S53876; MUID:95214615; PMID:7700229

A:Accession: S53876

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-149 <YAN>

A:Cross-references: GB:S77099; NID:9913977; PID:9913978

C:Genetics:

A:Gene: janA

A:Cross-references: FlyBase:FBgn0015151

A:Introns: 78/1; 119/3

Query Match	46.2%	Score 115;	DB 2;	Length 149;
Best Local Similarity	46.8%	Pred. No. 8.6e-08;		
Matches 21; Conservative	6;	Mismatches 16;	Indels 0;	Gaps 0;
DB	76	YHADIYDKVSGDMKQKGCDCGGRISHQSDKKIHVGYGS 43		
		: :: :: :		
		76 WHADYGRTOGINKGTGLDTGCTGGGRIGHMPGKKYLVGHHS 118		

RESULT 2

A32317 sex-regulated protein janA - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 29-Jan-1990 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000

C:Accession: A32317

R:Yanicosas, C.; Vincent, A.; Lepeant, J.A.

Mol. Cell. Biol. 9, 2526-2535, 1989

A:Title: Transcriptional and posttranscriptional regulation contributes to the sex-re

A:Reference number: A32317; MUID:89343970; PMID:2503707

A:Accession: A32317

A:Status: Preliminary

A:Molecule type: DNA; mRNA

A:Residues: 1-135 <YAN>

A:Cross-references: GB:M27033; NID:93522914; PIDN:AAC34203.1; PID:9387597

A:Note: the authors translated the codon TNG for residue 98 as Met

C:Genetics:

A:Gene: FlyBase:janA

A:Cross-references: FlyBase:FBgn0001280

Query Match	43.8%	Score 109;	DB 2;	Length 135;
Best Local Similarity	46.5%	Pred. No. 4.9e-07;		
Matches 20; Conservative	7;	Mismatches 16;	Indels 0;	Gaps 0;

```

Qy      1 YHADIYDKVSGMQKQGCDCCELTGGGRISQSDQKINHYGIS 43
        :|:::  :|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      62 WHADIFEREVEYFKKLTGLRADCPGGGRIEHNPEKKYLVYGIS 104

```

RESULT 3
B32317
B32317

sex-regulated protein janh - fruit fly (*Drosophila melanogaster*)
C/Species: *Drosophila melanogaster*
C/Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 21-Jul-2000
C/Accession: B32317
C/Ryanicostas, C.; Vincent, A.; Lepesant, J.A.
Mol. Cell. Biol. 9, 2526-2535, 1989
A/Title: Transcriptional and posttranscriptional regulation contributes to the sex-regulation of the sex-regulated protein janh
A/Reference number: A32317; MUID:89343970; PMID:2505707
A/Accession: B32317
A/Status: preliminary
A/Molecule type: DNA; mRNA
A/Residues: 1-140 <YAN>
A/Cross-references: GB:M27033; NID:g3522914; PIDN:AAC34204.1; PID:g387598.
C/Genetics:
A/Gene: FlyBase:janh
A/Cross-references: FlyBase:FBgn0001281

Query Match	92	Score	92	DB	2	Length	140
Best Local Similarity	35.7%	Pred. No.	8.7e-05				
Matches	15	Conservative	14	Mismatches	13	Indels	0
				Gaps	0		
QY	2	NADIDYKVGDMQKQGCDCCECGGGGRISHSOSDKRIHYVYGS	43				
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :					
Db	69	HLAVFDPSILTEELPEBICAKILGGGRILNEADKKRIKIGTIS	110				

RESULT 4
 T21833
 hypothetical protein F36A2.8 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21833
 R:Lennard, N.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19476
 B:Accession: T21833
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-115 <Wild>
 A:Cross-references: EMBL:Z81077; PIDN:CA803070.1; GSPDB:GN00019; CESP:F36A2.8
 A:Experimental source: clone F36A2
 C:Genetics:
 A:Gene: CESP:F36A2.8
 A:Map position: 1
 A:Introns: 22/3; 52/3

```
Query Match Similarity      31.5%; Score 78.5; DB 2; Length 115;
Best Local Similarity       39.5%; Pred. No.0.0044;
Matches    17; Conservative   4; Mismatches    21; Indels     1; Gaps     1;
```

Oy 1 YHADITYDKVSGDMQKCGCDCELCGGGRISHQSODKKIHNYGTS 43
 :
Db 43 FHDDILQETKRSSAPSD-LTKLCKVGGRIRHDEIGRDLIVYGS 84
 :
 :
 :
 :
 :
 :
 :
 :

RESULT 5

S53877 sex-regulated protein janus B - fruit fly (*Drosophila pseudoobscura*)
C:Species: Drosophila pseudoobscura
C>Date: 18-Oct-1995 #sequence revision 03-Nov-1995 #text_change 29-Jan-1999
C:Accession: S53877
R:Yanicostras, C.; Ferrer, P.; Vincent, A.; Lepesant, J.A.
Mol. Gen. Genet. 246, 549-560, 1995

A>Title: Separate cis-regulatory sequences control expression of serendipity beta and gamma

A:Reference number: S53876; MOID:95214615; PMID:7700229

A:Accession: S53877
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140 <FAN>
A:Cross-references: GB:S77099; NID:g913977; PID:g913979
C:Genetics:
A:Gene: janb
A:Cross-references: FlyBase:FBgn0014784
A:introns: 18/1; 70/1; 111/3

Query Match	30.9%	Score	11	DB	2	Length	140
Best Local Similarity	40.0%	Pred. No.	0.0083				
Matches	16	Conservative	7	Mismatches	17	Indels	0
						Gaps	0
QY	2	HADIDKVS	GDMOKOGDCECLGGGRISHQSDOKRHHYG	41			
Db	69	HDDIYYKNV	RAMGKLTCTKCLGGGDMNDGSKRKGIHG	108			

RESULT 6
S20806
swallow protein - fruit fly (*Drosophila melanogaster*)
C.Species: *Drosophila melanogaster*
C.Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 04-Mar-2000
C.Accession: S20806
R.Chao, Y.C.; Donahue, K.M.; Pokrywka, N.J.; Stephenson, E.C.
submitted to the EMBL Data Library, September 1990
A.Description: Sequence of swallow, a gene required for the localization of bicoid mesoderm in the fruit fly embryo.
A.Reference number: S20806
A.Accession: S20806
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-548 <CHA>
A.Cross-References: EMBL: X56023; NID: g8683; PID: g8684
C.Genetics:
A.Gene: FLYBase:swa
A.Cross-references: FLYBase: FBgn0003655
A.Introns: 91/1; 371/1
C.Superfamily: *Drosophila* swallow protein
C.Keywords: nucleus

Query Match	24.3%	Score 60.5	DB 2	Length 548
Best local Similarity	44.4%	Pred. No. 4.6		
Matches 16	Conservative 3	Mismatches 16	Indels 1	Gaps 1
QY	4	DIYDK-VSGDPMKQCGDCCECTGSGGRISRSHSQDKIKH	38	
Db	83	DDYDEVDGDDHRLGCEKAPLGSGRSSKAVSYDDIH	118	

```

RESULT 7
T33531
hypothetical protein F58E1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T33531
R:Wamsley, P.; Twyman, B.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid F58E1.
A:Reference number: Z21365
A:Accession: T33531
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-477 <MAN>
A:Cross-references: EMBL:AF098995; PIDB:MAC67482.1; GSPDB:GN00020; CESP:F58E1.4
A:Experimental source: strain Bristol NZ; clone F58E1
C:Genetics:
A:Gene: CESP:F58E1.4
A:Map position: 2
A:introns: 32/1: 104/1: 269/3: 424/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T05A6.4

Query Match      24.1%; Score 60; DB 2; Length 477;

```

Accession: S35191; #sequen

A; Molecule type: DNA
A; Residues: 1-296 <GAL>

Accession: S35191; #sequen

A; Molecule type: DNA
A; Residues: 1-296 <GAL>

C:Comment: The 100K protein is a late nonstructural protein involved in transport of hex
C:Genetics:
A:Map position: 66.4-71.4
C:Superfamily: adenovirus late 100K protein
C:Keywords: late protein

Query Match 22.7%; Score 56.5; DB 1; Length 296;
Best Local Similarity 50.0%; Pred. No. 8.5;
Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 1 YHADIDKVSQDMQKQGCDCCEC 22
DB 40 YHSDIMEDVSGDGLLE-CHCRC 60

RESULT 13

WMA015

late 100K protein - human adenovirus 5

C:Species: Mastadenovirus h5 (human adenovirus 5)

C>Date: 02-Apr-1992 #sequence_revision 30-Sep-1992 #text_change 05-Sep-1997

C:Accession: A39449; A03838

R:Chroboczek, J.; Bieher, F.; Jacrot, B.

Virology 186, 280-285, 1992

A>Title: The sequence of the genome of adenovirus type 5 and its comparison with the gen

A:Reference number: A39449; MUID:92087470; PMID:1727603

A:Accession: A39449

A:Molecule type: DNA

A:Residues: 1-807 <CHR>

A:Cross-references: GB:M73260

R:Knijfer, W.; van Schaik, F.M.A.; Sussenbach, J.S.

Nucleic Acids Res. 9, 4439-4457, 1981

A>Title: Structure and organization of the gene coding for the DNA binding protein of ad

A:Reference number: A39735; MUID:82059474; PMID:6117824

A:Accession: A03838

A:Molecule type: DNA

A:Residues: 1-593 <KRU>

A:Cross-references: GB:J01966; NID:9209789; PID:9209807

C:Comment: The 100K protein is a late nonstructural protein involved in transport of hex

C:Genetics:

A:Map position: 66.4-71.4

C:Superfamily: adenovirus late 100K protein

C:Keywords: late protein; nonstructural protein

Query Match 22.7%; Score 56.5; DB 1; Length 807;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

RESULT 14

A65215 hypothetical 49.4 kD protein in sox-r-acs intergenic region - Escherichia coli (strain K-

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: A65215

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A65215

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-430 <BLAT>

A:Cross-references: GB:AE000480; GB:U00096; NID:92367344; PIDN:AACT7036.1; PID:91790502;

C:Genetics:

A:Experimental source: strain K-12, substrain MG1655

A:Gene: yjcf

Query Match 22.3%; Score 55.5; DB 2; Length 430;

Best Local Similarity 35.0%; Pred. No. 17;
Matches 14; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

QY 1 YHADIDKVSQDMQKQGCDCCECLGGGRISHQSDKKIHY 40
DB 53 FKMDHEKVTGILQKMKRIYDLSIPSGKITKFT-DRIKHY 91

RESULT 15

S02532 tubulin beta-1 chain - slime mold (Physarum polycephalum) (fragment)

C:Species: Physarum polycephalum

C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 23-Feb-1997

C:Accession: S02532

R:Merenskiold, A.K.; Poetsch, B.; Haugli, F.

Eur. J. Biochem. 174, 491-495, 1988

A>Title: Cloning and expression of a beta tubulin gene of Physarum polycephalum.

A:Reference number: S02532; MUID:88271316; PMID:3391166

A:Accession: S02532

A:Molecule type: DNA

A:Cross-references: EMBL:X12371

C:Genetics:

A:Gene: betB1

A:Introns: 95/1

C:Superfamily: tubulin

C:Keywords: nucleus

Query Match 22.1%; Score 55; DB 2; Length 204;
Best Local Similarity 35.7%; Pred. No. 9.4;
Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 ADIDKVSQDMQKQGCDCCECLGGGRISH 30
DB 107 AELIDSVLDVVRKEAESCDLQGFQIAH 134

Search completed: May 7, 2003, 19:07:35
Job time: 11.4862 secs

1	119	47.8	145	1	JANA_DROPS	P54364	drosophila
2	109	43.8	135	1	JANA_DROME	P20348	drosophila
3	92	36.9	140	1	JANB_DROME	P20349	drosophila
4	83	33.3	140	1	JANB_DROPS	P54365	drosophila
5	60.5	24.3	548	1	SWA_DROME	P40688	drosophila
6	57	22.9	445	1	TBB1_TRTVI	P31864	trichoderm
7	56.5	22.7	805	1	L100_ADE02	P24932	human aden
8	56.5	22.5	807	1	L100_ADE05	P24933	human aden
9	56	22.5	444	1	TBB2_DNUCA	P24934	human aden
10	55.5	22.3	430	1	TBB2_DNUCI	P33704	daucus car
11	55	22.1	193	1	VANI_VIBAN	P33707	eschrichili
12	55	22.1	423	1	TBB2_PORPU	P74945	vibrio angu
13	55	22.1	444	1	TBB_ACHKL	P50260	porphyra pu
14	55	22.1	444	1	TBB_ACHKL	P20802	achlya klei
15	55	22.1	444	1	TBB_PHYCI	P05887	phytophtho
16	55	22.1	451	1	TBB_MNGRI	P34108	neogleria g
17	55	22.1	467	1	TBB1_PIRFO	P07436	physarum pu
18	54.5	21.9	259	1	HIS6_LACLA	P02133	lactococcus
19	54	21.7	441	1	TBB_BABBO	Q04709	babesia bov
20	54	21.7	442	1	TBB1_PARTE	P33188	paramecium
21	54	21.7	442	1	TBB_STYLE	P11857	stylyonychia
22	54	21.7	443	1	TBB1_ANEPH	P33630	anemia phy
23	54	21.7	443	1	TBB_TERRY	P10876	tetrahymen
24	54	21.7	444	1	TBB_TERTH	P1352	tetrahymen
25	54	21.7	444	1	TBB_EUPFO	Q9n2n6	euplotes fo
26	54	21.7	445	1	TBB1_SOYBN	Q08115	euplotes oc
27	54	21.7	445	1	TBB_LEIME	P14559	glycine max
28	54	21.7	445	1	TBB_PLAFA	P21148	leishmania
29	54	21.7	446	1	TBB_EUPCR	P14140	plasmodium
30	54	21.7	447	1	TBB1_CYAPA	P20365	euplotes cr
31	54	21.7	449	1	TBB_TOXGO	Q9sw41	cyanoophora
32	54	21.7	449	1	TBB4_PORPU	P10878	toxoplasma
33	53.5	21.5	162	1	BAR2_CHIPA	P50262	porphyra pu
						P08725	chironomus

34	53.5	1603	1	PSC_DROME	P35820	drosophila
35	21.5	341	1	TBB_HALDI	P41386	halictus di
36	53	366	1	TBB1_AVESA	P25862	aveana sativ
37	21.3	408	1	TBB3_SOYBN	P28551	glycine max
38	53	411	1	TBB2_ANEPH	P33631	anemaphys
39	53	440	1	TBB3_PEA	P29502	plum sativ
40	21.3	443	1	TBB2_XENLA	P33602	xenopus lae
41	53	444	1	TBB1_HUMAN	P07437	homo sapien
42	53	21.3	1	TBB1_ORYSA	P37832	oryza sativ
43	53	21.3	1	TBB2_MAIZE	P18026	zea mays (m
44	53	444	1	TBB7_CHICK	P09244	gallus galli
45	21.3	444	1	TBB9_ARARH	P29517	arabidopsist

ALIGNMENTS

RESULT 1
JANA DROPS

ID	JANA_DROPS	STANDARD;	PRT;	146 AA
AC	P5A36A:			

DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Fact occurrence updated)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sex-regulated protein in fatuc

GN
JANA.
Drosophila pseudoobscura (Fruit

OC Eukaryota; Metazoa; Arthropoda; Mand
OC Insecta: Pterivota: Hymenoptera: Endent

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila
OX NCBI TaxID=7237.

RN (1) BP SEQUENCE FROM A

RA MEDLINE=95214615; PubMed=77002229
Yanicostas C, Ferrar P, Vincent

RT "Separate cis-regulatory sequences control expression of serendipity
RT beta and janus A: two immediately adjacent *Drosophila* genes"

CC -1- SIMILARITY: WITH SEX-REGULATED
RL Mol. Genet. 246:549-560 (1995)

cc This SWISS-PROT entry is copyright. It is produced
cc

cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
cc the European Bioinformatics Institute. There are no restrictions on use
cc

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```
cc entires requires a license agreement (
cc or send an email to license@isb-sib.ch)
```

CC
DR EMBL; S77099; AAB33912.2; -

DN
lyase; rbgnuvici1; upsej ana.
SEQUENCE 146 AA; 16291 MW; 02F7DEC570070D7B CRC64
50

Query Match 47.8%; Score 119; DB 1; Length 146;

```

Best local similarity  51.2%;  Pred. NO: 3e-09;
Matches  21;  Conservative  7;  Mismatches  13;  Indels  0;  Gaps  0

```

0Y 3 ADIYDKVSGDMQKQGCDECELGGRISHQSQDKIHVYGS 43

Db 75 ADIYERTQGTIKGTGLDTECLGGRIEHNPEKKYLKVGHS 115

JANA_DROME

P20348; Q9VAB6;

		DT	01-NOV-1995 (Rel. 32, Last sequence update)
--	--	----	---

Sex-regulated protein janus-A.

Drosophila melanogaster (Fruit fly).

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of Drosophila melanogaster."
 CC Science 287:2185-2195(2000).
 CC -1- TISSUE SPECIFICITY: GERM LINE CELLS OF ADULT MALES.
 CC -1- DEVELOPMENTAL STAGE: FROM THE THIRD-INSTAR LARVAL STAGE TO THE
 CC ADULT STAGE.
 CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-A.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M27033; AAC34204.1; -
 DR EMBL: AE003772; AAF56996.1; -
 DR PIR: B32317; B32317;
 DR Flybase: FBgn0001281; JanB.
 SQ SEQUENCE 140 AA: 15859 MW: 5750CC28BB61D7FE CRC64;
 Query Match 36.9%; Score 92; DB 1; Length 140;
 Best Local Similarity 35.7%; Pred. No. 1.4e-05;
 Matches 15; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
 QY 2 HADIDKVSQDMKOGCCECLGGGRISHQSDKKIHVYGS 43
 DB 69 HLAIVEDSILELEPEGICAKILGGRIINAEAKRKIKIYGT 110
 RESULT 4
 ID JANB_DROPS STANDARD; PRT; 140 AA.
 AC P54365;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sex-regulated protein Janus-B.
 GN JanB.
 OS Drosophila pseudoobscura (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CC NCBI_Taxid=7237;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95214615; PubMed=7700229;
 RA Yancostas C., Ferrer P., Vincent A., Lepesant J.-A.;
 RT "Separate cis-regulatory sequences control expression of serendipity
 RL beta and janus A, two immediately adjacent Drosophila genes.";
 RL Mol. Genet. 246:549-560(1995).
 CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-A.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S77099; AAB33913.2; -
 DR Flybase: FBgn0014784; Dpse\JanB.
 SQ SEQUENCE 140 AA: 15812 MW: 01B87FE1F0AEF53A CRC64;
 Query Match 33.3%; Score 83; DB 1; Length 140;
 Best Local Similarity 40.0%; Pred. No. 0.00024;
 Matches 16; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
 QY 2 HADIDKVSQDMKOGCCECLGGGRISHQSDKKIHVYGS 41
 DB 69 HDDIYKNNVRAEMKLGCTKCI GGGKMDNESARKIEIHG 108

RESULT 5
 ID SWA_DROME STANDARD; PRT; 548 AA.
 AC P40688;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Swallow protein.
 GN SWA.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=92217243; PubMed=1806330;
 RA Chao Y.-C., Donahue K.M., Pokrywka N.J., Stephenson E.C.;
 RT "Sequence of swallow, a gene required for the localization of bicoid
 RL message in Drosophila eggs.";
 RL Dev. Genet. 12:333-341(1991).
 RN [2]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=94116446; PubMed=7507030;
 RA Hegde J., Stephenson E.C.;
 RT "Distribution of swallow protein in egg chambers and embryos of
 RL Drosophila melanogaster.";
 RL Development 119:457-470(1993).
 CC -1- FUNCTION: HAS A ROLE IN LOCALIZING BICOID MRNA AT THE ANTERIOR
 CC MARGIN OF THE OOCYTE DURING OOGENESIS, AND A POORLY CHARACTERIZED
 CC ROLE IN NUCLEAR DIVISIONS IN EARLY EMBRYOGENESIS.
 CC -1- SUBUNIT: MAY BE CONSTITUTED OF AN HOMO- OR HETERODIMER.
 CC -1- SUBCELLULAR LOCATION: UNIFORMLY DISTRIBUTED IN EGGS, BECOMES
 CC LOCALIZED TO THE NUCLEI DURING EARLY MITOTIC DIVISIONS IN EARLY
 CC EMBRYOGENESIS. SWALLOW ENTERS EACH NUCLEUS AT THE BEGINNING OF
 CC MITOSIS, OCCUPIES A POSITION COMPLEMENTARY TO THAT OF CONDENSED
 CC CHROMATIN, AND LEAVES EACH NUCLEUS AT THE END OF MITOSIS.
 CC -1- TISSUE SPECIFICITY: FOUND IN EARLY EMBRYOS AND IN OVARIES. NOT
 CC DETECTABLE IN MALES OR IN FEMALE BODIES EXCLUDING OVARIES.
 CC -1- DEVELOPMENTAL STAGE: PRESENT THROUGHOUT OOGENESIS AND AFTER
 CC FERTILIZATION. PRESENT ONLY IN 0-6 HOUR EMBRYOS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X56023; CA39500.1; -
 DR PIR: S20806;
 DR Flybase: FBgn0003655; swa.
 DR InterPro: IPR000504; RNA_rec.mot.
 KW Developmental protein; Nuclear protein; Mitosis.
 SQ SEQUENCE 548 AA: 62070 MW: B8E0351C7B005159 CRC64;
 Query Match 24.3%; Score 60.5; DB 1; Length 548;
 Best Local Similarity 44.4%; Pred. No. 1.1;
 Matches 16; Conservative 3; Mismatches 16; Indels 1; Gaps 1;
 QY 4 DIYDK-VSGDMKOGCCECLGGGRISHQSDKKIH 38
 DB 83 DDYEDVDGDHDKLGCEKAPLGSGRSKAVSYDITH 118
 RESULT 6
 ID TBBL_TRIVI STANDARD; PRT; 446 AA.
 AC P31864;

DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1993 (Rel. 38, Last annotation update)
 DE Tubulin beta-1 chain.
 GN TUB1.
 OS Trichoderma viride.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Trichoderma.
 OX NCBI_TaxID=5547;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=T9 BR47;
 RX MEDLINE=93341462; PubMed=8341264;
 RA Goldman G.H., Temmerman W., Herrera-Estrella A., Jacobs D.,
 RA Contreras R., van Montagu M., "A nucleotide substitution in one of the beta-tubulin genes of
 RT Trichoderma viride confers resistance to the antimicrobial drug methyl
 RT benzimidazole-2-yl-carbamate."
 RL Mol. Gen. Genet. 240:73-80(1993).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: Z15054; CAA78764.1; -
 CC PIR: S25553; S25553.
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PR01161; TUBULIN.
 DR PROSITE: PS00227; TUBULIN.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 KW Microtubules; GTP-binding.
 FT NP_BIND 142 148
 FT SEQUENCE 446 AA; 49625 MW; B43E7B478B1CE13 CRC64;
 SQ
 QY 3 ADIYKVSQDMQKQDCCEIAGGGRISH 30
 Db 112 AELVDQVLDVVRREANGCEELQGFQITH 139
 Query Match 22.9%; Score 57; DB 1; Length 446;
 Best Local Similarity 35.7%; Pred. No. 2.7;
 Matches 10; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

RP SEQUENCE OF 510-805 FROM N.A.
 RX MEDLINE=80004828; PubMed=478297;
 RA Galibert F., Herisse J., Courtois G.;
 RT "Nucleotide sequence of the EORI-F fragment of adenovirus 2 genome."
 RL Gene 6:1-22(1979).
 CC -1- FUNCTION: THE 100 KDA PROTEIN IS A LATE NONSTRUCTURAL PROTEIN
 CC INVOLVED IN TRANSPORT OF HEXON FROM CYTOPLASM TO THE NUCLEUS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: J01917; AAA92218.1; -
 CC PIR: A03838; WMADL2.
 DR InterPro: IPR003381; Adeno_100.
 DR Pfam: PF02438; adeno_100; 1.
 KW Late protein; Transport.
 SQ SEQUENCE 805 AA; 90137 MW; 4C772CF19CA4C573 CRC64;
 Query Match 22.7%; Score 56.5; DB 1; Length 805;
 Best Local Similarity 50.0%; Pred. No. 5.8;
 Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
 QY 1 YHADYKVSQDMQKQDCCEC 22
 Db 549 YHSDIMEDVSGDGLLE-CHCRC 569
 RESULT 8
 ID L100_ADE05 STANDARD; PRT; 807 AA.
 AC P24933; P03268;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DE Late 100 kDa protein.
 DE Human adenovirus type 5.
 OS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OC NCBI_TaxID=28285;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92087470; PubMed=1727603;
 RA Chroboczek J., Bieher F., Jacrot B.;
 RT "The sequence of the genome of adenovirus type 5 and its comparison
 RT with the genome of adenovirus type 2."
 RL Virology 186:280-285(1992).
 RN (2)
 RP SEQUENCE OF 1-593 FROM N.A.
 RX MEDLINE=82059474; PubMed=6117824;
 RA Kruljer W., van Schaik F.M.A., Sussenbach J.S.;
 RT "Structure and organization of the gene coding for the DNA binding
 RT protein of adenovirus type 5."
 RL Nucleic Acids Res. 9:4439-4457(1981).
 CC -1- FUNCTION: THE 100 KDA PROTEIN IS A LATE NONSTRUCTURAL PROTEIN
 CC INVOLVED IN TRANSPORT OF HEXON FROM CYTOPLASM TO THE NUCLEUS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: J01966; AAD15296.1; -
 CC EMBL: M73260; AAA96412.1; ALT_SEQ.
 DR EMBL: X02997; CAA26756.1; -
 DR PIR: A39449; WMAD15.
 DR InterPro: IPR003381; Adeno_100.

DR Pfam: PF02438; adeno_100; 1.
 KW Late protein; Transport.
 SQ SEQUENCE 807 AA; 90212 MW; CF17561EDA7CBADF CRC64;

Query Match 22.7%; Score 56.5; DB 1; Length 807;
 Best Local Similarity 50.0%; Pred. No. 5.8;
 Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OY 1 YHADIYKVSQDMQKGCDCRC 22
 DB 549 YHSDIMEDVSGDGLLE-CHCRC 569

RESULT 9

TBB2_DAUCA STANDARD; PRT; 444 AA.
 AC Q39697;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tubulin beta-2 chain (Beta-2 tubulin).
 GN TUBB2.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Kintokai;
 RX MEDLINE=9738494; PubMed=9195054;
 RA Okamura S., Naito K., Sonehara S., Ohkawa H., Kuramori S., Tatsuta M.,
 RT Minamizono M., Katsuka T.;
 RT "Characterization of the carrot beta-tubulin gene coding a divergent
 RT isotype, beta-2.";
 RL Cell Struct. Funct. 22:291-298(1997).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- TISSUE SPECIFICITY: FOUND IN AREAS OF RAPIDLY DIVIDING TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation;
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U63927; AAB64308.1;
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR Pfam: PF00091; tubulin_F1sz.
 DR PRINTS: PR01161; TUBULIN.1.
 DR PROSITE: PS00327; TUBULIN.1.
 DR PROSITE: PS00228; TUBULIN_B/AUTOREG.1.
 KW Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 140 146 GTP (POTENTIAL).
 SQ SEQUENCE 444 AA; 49571 MW; 359D3C7433DB1E88 CRC64;

Query Match 22.5%; Score 56; DB 1; Length 444;
 Best Local Similarity 35.7%; Pred. No. 3.7;
 Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

OY 3 ADIYKVSQDMQKGCDCRCGLGGGRISH 30
 DB 110 AELIDSVLDVVRKKAENCCGLGGPOVCH 137

RESULT 10

YJCF_ECOLI
 ID YJCF_ECOLI STANDARD; PRT; 430 AA.
 AC P32704;

DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yjcf.
 GN YJCF OR B4066.
 OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

OX NCBI_TaxID=562;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;

RX MEDLINE=94089392; PubMed=8265357;
 RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,

RT Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the

RT region from 89.2 to 92.8 minutes.";
 RL Nucleic Acids Res. 21:5408-5417(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation;
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: U00006; AAC43160.1;
 DR EMBL: AE000480; AAC77036.1;

DR Ecocore: E01941; YJCF.
 DR InterPro: IPR001646; 5peptide_repeat.

DR Pfam: PF08085; pentapeptide_4.
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 430 AA; 49378 MW; C4591C44F256F9C5 CRC64;

Query Match 22.3%; Score 55.5; DB 1; Length 430;
 Best Local Similarity 35.0%; Pred. No. 4.2;
 Matches 14; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

OY 1 YHADIYKVSQDMQKGCDCRCGLGGGRISHOSODKKIHV 40
 DB 53 FKMDDHEKVIYGLIKMKRKYDSLPGKITKET-DRKIKY 91

RESULT 11

VANT_VIBAN STANDARD; PRT; 193 AA.
 ID VANT_VIBAN

AC P74945;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Autoinducer synthesis protein vanti.

GN VANT.

OS Vibrio anguillarum (Listonella anguillarum).
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Listonella.

OX NCBI_TaxID=55601;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=NB10 / Serotype O1;

RX MEDLINE=97284506; PubMed=9139920;
 RA Milton D.L., Hardman A., Camara M., Chhabra S.R., Bycroft B.W.,

RT Stewart G.S.A.B., Williams P.;
 RT "Quorum sensing in Vibrio anguillarum: characterization of the

RT vanti/vank locus and identification of the autoinducer N-(3-
 RT oxodecanoyl)-L-homoserine lactone.";

RL J. Bacteriol. 179:3004-3012(1997).
 CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF N-(3-OXODECANOYL)-L-

CC HOMOSERINE LACTONE (ODHL), AN AUTOINDUCER MOLECULE WHICH BINDS TO
 CC VANT.
 CC -1- SIMILARITY: BELONGS TO THE AUTOINDUCER SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U69677; AAC45212.1; -
 DR InterPro: IPR001690; Autoind_synth.
 DR Pfam: PF00765; Autoind_synth; 1.
 DR PRINTS: PR01549; AUTOINDCRSYN.
 DR ProDom: PD002752; Autoind_synth; 1.
 DR PROSITE: PS00949; AUTOINDUCERS_SYTH; 1.
 DR Quorum sensing; Autoinducer synthesis.
 KW SEQUENCE 193 AA; 22127 MW; 238E760CD8701188 CRC64;

Query Match 22.1%; Score 55; DB 1; Length 193;
 Best Local Similarity 30.4%; Pred. No. 2.2; Mismatches 14; Gaps 2;
 Matches 14; Conservative 7; Indels 14; Gaps 2;

QY 1 YHADYKVSQD-----MOKGDCCELCGGRIHQSDQKTHVYG 41
 DB 136 YHNAVYVTSASVEKIKRMGIPCERLG-----DKVHLIG 172

RESULT 12
 TB22_PORPU STANDARD; PRT; 423 AA.
 AC P50260;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Tubulin beta-2 chain (Fragment).
 GN TUBB2.
 OS Porphyra purpurea.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
 OX NCBI_TaxID=2787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Avonport;
 RA Mackay R.M., Gallant J.W.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z67992; CAA91940.1; -
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; PARTIAL.
 KW Microtubules; GTP-binding; Multigene family.
 FT NON_TER 1
 FT NP_BIND 115 121 GTP (POTENTIAL).
 SQ SEQUENCE 423 AA; 47367 MW; 4EB030756DD3A45A CRC64;

Query Match 22.1%; Score 55; DB 1; Length 423;
 Best Local Similarity 35.7%; Pred. No. 4.9; Mismatches 10; Indels 0; Gaps 0;
 Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 ADYDKVSGDMOKGDCCELCGGRIHQSDQKTHVYG 30
 DB 85 AELIDSVLDVYKRAESCDLCGGRIHQSDQKTHVYG 112

RESULT 13
 TB22_PORPU STANDARD; PRT; 444 AA.
 AC P20802;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Tubulin beta chain.
 DE Achlya klebsiana.
 OS Achlya klebsiana.
 CC Eukaryota; stramenopiles; Oomycetes; Saprolegniales; Saprolegniaceae;
 OC Achlya.
 OX NCBI_TaxID=4767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90368709; PubMed=2394720;
 RA Cameron L.E., Hutts J.-A., Thorlacius L., Lejohn H.B.;
 RL J. Biol. Chem. 265:15245-15252(1990).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DEVELOPMENTAL STAGE: SPORANGIUM FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: J05597; AAA63161.1; -
 DR PIR: A35885; A35885.
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PR01161; TUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 KW Microtubules; GTP-binding.
 FT NP_BIND 139 145 GTP (POTENTIAL).
 SQ SEQUENCE 444 AA; 49851 MW; 7B3B25B29626B5E CRC64;

Query Match 22.1%; Score 55; DB 1; Length 444;
 Best Local Similarity 35.7%; Pred. No. 5.1; Mismatches 10; Indels 0; Gaps 0;
 Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 ADYDKVSGDMOKGDCCELCGGRIHQSDQKTHVYG 30
 DB 109 AELIDSVLDVYKRAESCDLCGGRIHQSDQKTHVYG 136

RESULT 14
 TB22_PORPU STANDARD; PRT; 444 AA.
 AC P20802;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Tubulin beta chain.
 DE Achlya klebsiana.
 OS Achlya klebsiana.
 CC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Pythiophthora.

```

OX NCB1_TaxID=4785;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DAR 52646;
RA Meerakoon N.D., Roberts J.K., Lehnen L.P. Jr., Wilkinson J.M.,
RT Marshall J.S., Hardham A.R.;
RT "Isolation and characterization of the single beta-tubulin gene in
RT Phytophthora cinnamomi.";
RL Mycologia 90:85-95(1998).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U22050; AAC05441.1;
DR InterPro: IPR002453; Beta_tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin_1.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
KW Microtubules; GTP-binding.
FT NP_BIND 140 146 GTP (POTENTIAL).
SQ SEQUENCE 444 AA; 49922 MW; 8392C8C288378929 CRC64;

Query Match 22.1%; Score 55; DB 1; Length 444;
Best Local Similarity 35.7%; Pred. No. 5.1;
Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 ADIYKVSQDMKRGCDCECLGGGRISH 30
DB 110 AELIDSVLDVVRKEASCDCLQGFIATH 137

RESULT 15
TBB_NAEGR
ID TBB_NAEGR STANDARD; PRT; 451 AA.
AC P34108;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta chain.
OS Naegleria gruberi.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX NCB1_TaxID=5762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEG;
RX MEDLINE=94118284; Pubmed=8289261;
RA Lai E.Y., Remillard S.P., Fulton C.;
RT "A beta-tubulin gene of Naegleria encodes a carboxy-terminal
RT lysine. Aromatic amino acids are conserved at carboxy terminl.";
RL J. Mol. Biol. 235:377-388(1994).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

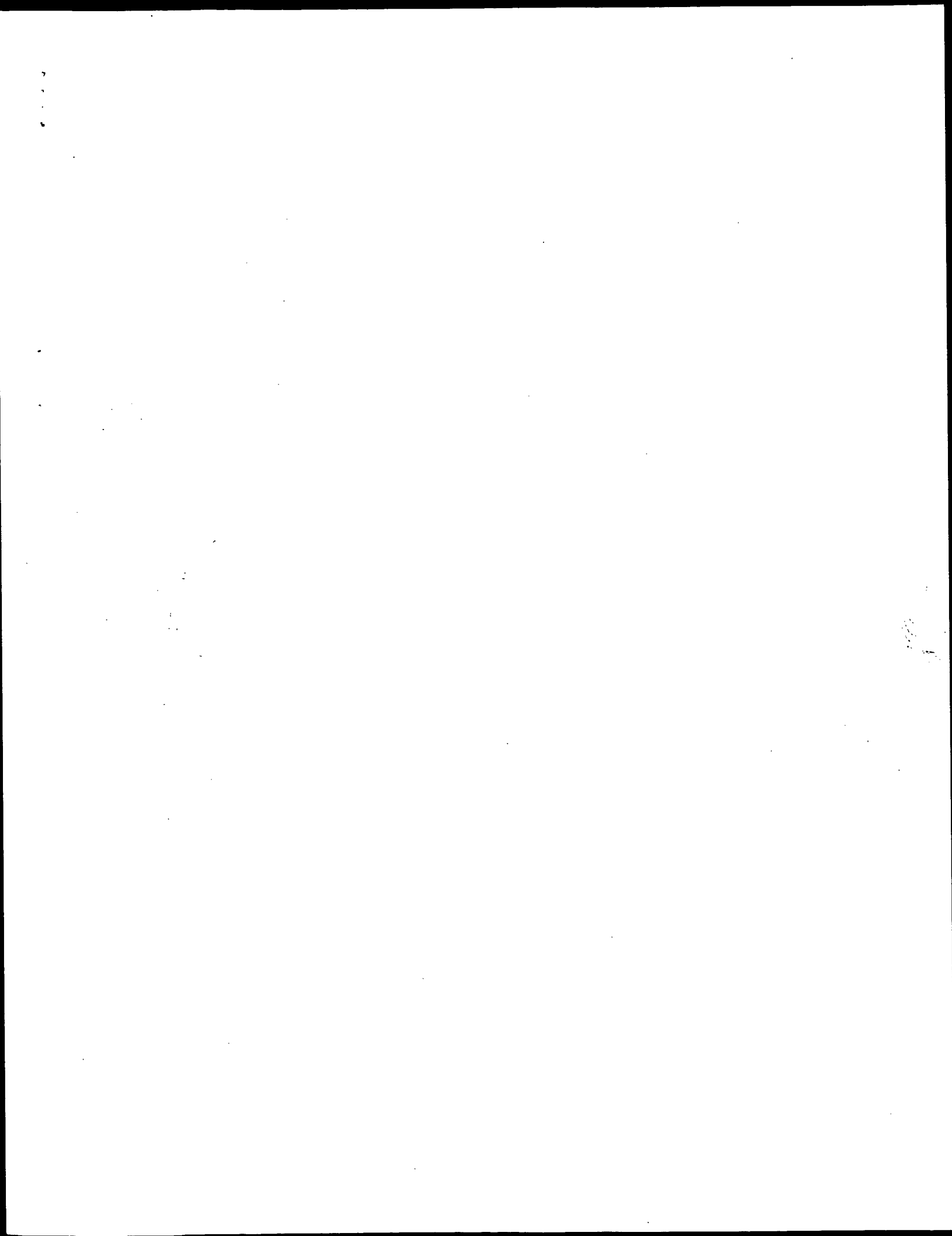
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z13961; CAAT78362.1;
DR PIR: S30514; S30514.
DR InterPro: IPR002453; Beta_tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin_1.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
KW Microtubules; GTP-binding.
FT NP_BIND 140 146 GTP (POTENTIAL).
SQ SEQUENCE 451 AA; 50511 MW; 842777EC706BDC1 CRC64;

Query Match 22.1%; Score 55; DB 1; Length 451;
Best Local Similarity 35.7%; Pred. No. 5.2;
Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 ADIYKVSQDMKRGCDCECLGGGRISH 30
DB 110 AELIDSVLDVVRKEASCDCLQGFIATH 137

```

Search completed: May 7, 2003, 19:04:02
Job time : 6.24771 secs



GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:00:25 ; Search time 30.4771 Seconds
(without alignments)
297.472 Million cell updates/sec

Title: US-09-914-831-5

Sequence: 1 YIADYDKVSGDMKQGCDC.....GGRIHQSDKKIHVGYSM 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:
2: SP-bacteria:
3: SP-fungi:
4: SP-human:
5: SP-invertebrate:
6: SP-mammal:
7: SP-mhc:
8: SP-organelle:
9: SP-phage:
10: SP-plant:
11: SP-rodent:
12: SP-virus:
13: SP-vertebrate:
14: SP-unclassified:
15: SP-virus:
16: SP-bacteriophage:
17: SP-archaeop:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249	100.0	125	4	Q9NRX4
2	244	98.0	125	4	Q9H0Y3
3	220	88.4	124	11	Q9DAK9
4	113	45.4	135	5	Q9BH48
5	111	44.6	135	5	Q9BM97
6	111	44.6	135	5	Q9BM96
7	108	43.4	135	5	Q9BM98
8	107	43.0	102	5	Q9BH67
9	107	43.0	102	5	Q9BH65
10	107	43.0	104	5	Q9BMZ5
11	107	43.0	104	5	Q9BMZ1
12	107	43.0	109	5	Q9BMZ7
13	107	43.0	119	5	Q9BT45
14	107	43.0	135	5	Q9BM99
15	107	43.0	135	5	Q9SVY7
16	94	37.8	140	5	Q9BM82

17	93	37.3	140	5	Q9BM91	Q9BM91 drosophila
18	91	36.5	140	5	Q9BM90	Q9BM90 drosophila
19	91	36.5	140	5	Q9BM89	Q9BM89 drosophila
20	90	36.1	140	5	Q9BM92	Q9BM92 drosophila
21	90	36.1	140	5	Q9YL60	Q9YL60 drosophila
22	89	35.7	140	5	Q9BM94	Q9BM94 drosophila
23	89	35.7	140	5	Q9BM93	Q9BM93 drosophila
24	89	35.7	140	5	Q9BM85	Q9BM85 drosophila
25	89	35.7	140	5	Q9BM84	Q9BM84 drosophila
26	89	35.7	140	5	Q9BM95	Q9BM95 drosophila
27	88	35.3	140	5	Q9BM87	Q9BM87 drosophila
28	88	35.3	140	5	Q9BM86	Q9BM86 drosophila
29	88	35.3	140	5	Q9BM88	Q9BM88 drosophila
30	88	35.3	148	5	Q9Y170	Q9Y170 drosophila
31	87	34.9	140	5	Q9BM83	Q9BM83 drosophila
32	78.5	31.5	115	5	P90861	Q90861 caenorhabditis
33	63.5	25.5	276	13	Q9P059	Q9P059 brachydanio
34	61	24.5	487	2	Q9F901	Q9F901 streptococc
35	60.5	24.3	518	5	Q9W400	Q9W400 drosophila
36	60	24.1	266	10	Q9ALQ2	Q9ALQ2 oryza sativ
37	60	24.1	477	5	Q9TZG7	Q9TZG7 caenorhabditis
38	60	24.1	506	16	Q9P0C2	Q9P0C2 ureaplasma
39	58.5	23.5	382	16	Q9PME9	Q9PME9 campylobact
40	58	23.3	387	5	Q9BKC8	Q9BKC8 jakoba inca
41	58	23.3	387	10	Q9ZTL5	Q9ZTL5 guillardia
42	58	23.3	387	10	Q9S0J5	Q9S0J5 guillardia
43	58	23.3	566	10	Q9LKA4	Q9LKA4 arabidopsis
44	58	23.3	822	10	Q93XX2	Q93XX2 arabidopsis
45	58	23.3	846	10	Q9S588	Q9S588 arabidopsis

ALIGNMENTS

RESULT 1
ID Q9NRX4 PRELIMINARY; PRT; 125 AA.
AC Q9NRX4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sex-regulated protein janus-A (Cgi-202) (Similar to HSPC141 protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADRENAL GLAND;
RA Li Y., Shi J., Huang C., Jiang C., Ren S., Zhou J., Yu Y., Xu S., Wang Y., Fu G., Chen Z., Han Z.;
RA "A novel gene expressed in human adrenal gland."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RC TISSUE=ADRENAL GLAND;
RA Peng Y., Gu Y., Li Y., Fu S., Gu J., Gu W., Jiang C., Yu Y., Han Z., Wang Y., Chen Z., Fu G.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lin W.-C.;
RT "Identification of novel human genes by comparative proteomics."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF164795; AAF80759.1; -
DR EMBL: AF285119; AAC0156.1; -
DR EMBL: BC024648; AAH24648.1; -

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janusa.
 GN JANA.
 OS Drosophila erecta (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Melkleyohn C.D., Hauscheck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
 RT melanogaster Species Subgroup."
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL: AY013343; AAG50364.1;
 DR FLYBase: FBgn0043709; Dere\Jana.
 SQ SEQUENCE 135 AA; 15180 MW; DB294031FEDFD07C CRC64;

Query Match 44.6%; Score 111; DB 5; Length 135;
 Best Local Similarity 48.8%; Pred. No. 8.3e-08;
 Matches 21; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIYDKVSGDMOKGDCCECLGGGRISHQSODKKIHYGYS 43
 :||||: :| | | | | | : : | | | |
 Db 62 WHADIFREEEVFKKLGRLAECPCGGRIEHNPDKKYLYGYS 104

RESULT 6

ID 09BM96 PRELIMINARY; PRT; 135 AA.
 AC 09BM96;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janusa.
 GN JANA.
 OS Drosophila oreana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Melkleyohn C.D., Hauscheck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
 RT melanogaster Species Subgroup."
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL: AY013344; AAG50365.1;
 DR FLYBase: FBgn0043667; Dore\Jana.
 SQ SEQUENCE 135 AA; 15146 MW; DB21E237BCD1307C CRC64;

Query Match 44.6%; Score 111; DB 5; Length 135;
 Best Local Similarity 48.8%; Pred. No. 8.3e-08;
 Matches 21; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIYDKVSGDMOKGDCCECLGGGRISHQSODKKIHYGYS 43
 :||||: :| | | | | | : : | | | |
 Db 62 WHADIFREEEVFKKLGRLAECPCGGRIEHNPDKKYLYGYS 104

RESULT 7

ID 09BM98 PRELIMINARY; PRT; 135 AA.
 AC 09BM98;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Janusa.
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21219192; PubMed=11319264;
 RA Parsch J., Melkleyohn C.D., Hauscheck-Jungen E., Hunziker P.,
 RA Hartl D.L.;
 RT "Molecular Evolution of the ocnus and janus Genes in the Drosophila
 RT melanogaster Species Subgroup."
 RL Mol. Biol. Evol. 18:801-811(2001).
 DR EMBL: AY013340; AAG50361.1;
 DR FLYBase: FBgn0043691; Dmau\Jana.
 SQ SEQUENCE 135 AA; 15158 MW; 2F91E387BD70C1CC CRC64;

Query Match 43.4%; Score 108; DB 5; Length 135;
 Best Local Similarity 46.5%; Pred. No. 2.2e-07;
 Matches 20; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIYDKVSGDMOKGDCCECLGGGRISHQSODKKIHYGYS 43
 :||||: :| | | | | | : : | | | |
 Db 62 WHADIFREEEVFKKLGRLAECPCGGRIEHNPDKKYLYGYS 104

RESULT 8

ID 09BH68 PRELIMINARY; PRT; 102 AA.
 AC 09BH68;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KENYA.5; AND KENYA.12;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Kilian R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 RA Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species."
 RL Genetics 156:1913-1931(2000).
 DR EMBL: AF284454; AAG43468.1;
 DR EMBL: AF284455; AAG43470.1;
 DR FLYBase: FBgn0016348; Dsim\Jana.
 FT NON-TER
 SQ SEQUENCE 102 AA; 11608 MW; 7C1AC80B7E9369EB CRC64;

Query Match 43.0%; Score 107; DB 5; Length 102;
 Best Local Similarity 46.5%; Pred. No. 2.2e-07;
 Matches 20; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIYDKVSGDMOKGDCCECLGGGRISHQSODKKIHYGYS 43
 :||||: :| | | | | | : : | | | |
 Db 29 WHADIFREEEVFKKLGRLAECPCGGRIEHNPDKKYLYGYS 71

RESULT 9

ID 09BH67 PRELIMINARY; PRT; 102 AA.
 AC 09BH67;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Janus A (Fragment).
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CAP MALHEUREUX, AND PORT-LOUIS;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species.";
 RL Genetics 156:1913-1931(2000).
 DR EMBL; AF284457; AAC49474.1; -;
 DR EMBL; AF284458; AAC49476.1; -;
 DR FlyBase; FBgn0043691; Dmanu\Jana.
 FT NON_TER 1
 SQ SEQUENCE 102 AA; 11608 MW; 7C1AC80B7E9369EB CRC64;

Query Match 43.0%; Score 107; DB 5; Length 102;
 Best Local Similarity 46.5%; Pred. No. 2.2e-07;
 Matches 20; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIDKVSQDMKOGCCDCECLGGGRISHQSODKKIHVYGS 43
 DB 29 WHADIFEREVEYFKKLGARACPGGRIEHNPKKYLKLVYGS 71

RESULT 10

O9BMZ5 PRELIMINARY; PRT; 104 AA.
 AC O9BMZ5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila mauritiana (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7226;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1631;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species.";
 RL Genetics 156:1913-1931(2000).
 DR EMBL; AF284456; AAC49472.1; -;
 DR FlyBase; FBgn0043691; Dmanu\Jana.
 FT NON_TER 1
 SQ SEQUENCE 104 AA; 11834 MW; 690E994E3A961437 CRC64;

Query Match 43.0%; Score 107; DB 5; Length 104;
 Best Local Similarity 46.5%; Pred. No. 2.3e-07;
 Matches 20; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIDKVSQDMKOGCCDCECLGGGRISHQSODKKIHVYGS 43
 DB 31 WHADIFEREVEYFKKLGARACPGGRIEHNPKKYLKLVYGS 73

RESULT 11

O9BMZ1 PRELIMINARY; PRT; 104 AA.
 AC O9BMZ1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila sechellia (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7238;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=228;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species.";
 RL Genetics 156:1913-1931(2000).
 DR EMBL; AF284459; AAC49478.1; -;
 DR FlyBase; FBgn0043637; Dsec\Jana.
 FT NON_TER 1
 SQ SEQUENCE 104 AA; 11735 MW; 73BF32548B261427 CRC64;

Query Match 43.0%; Score 107; DB 5; Length 104;
 Best Local Similarity 46.5%; Pred. No. 2.3e-07;
 Matches 20; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIDKVSQDMKOGCCDCECLGGGRISHQSODKKIHVYGS 43
 DB 31 WHADIFEREVEYFKKLGARACPGGRIEHNPKKYLKLVYGS 73

RESULT 12

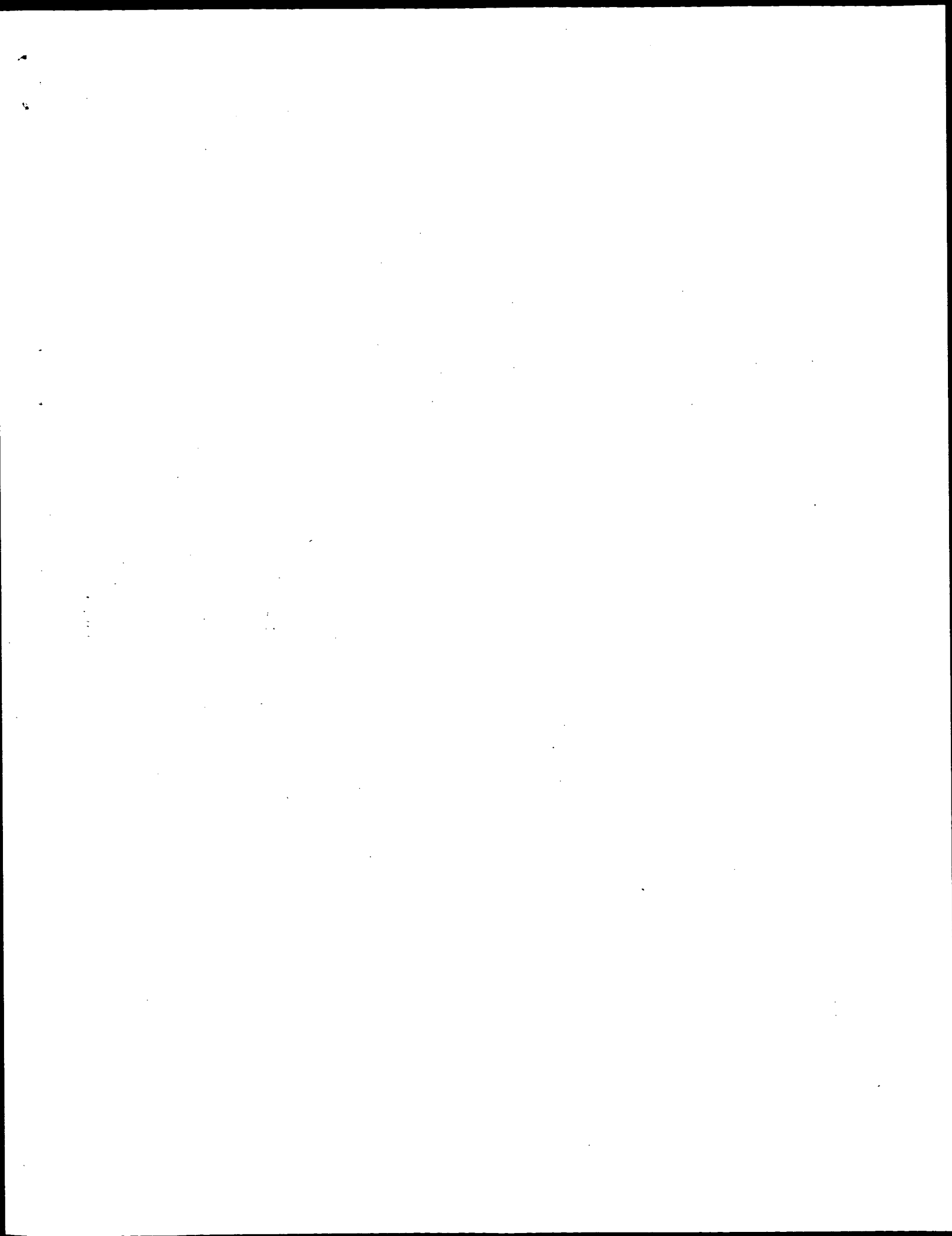
O9BMZ7 PRELIMINARY; PRT; 109 AA.
 AC O9BMZ7;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Janus A (Fragment).
 GN JANA.
 OS Drosophila simulans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KENYA_2;
 RX MEDLINE=20556156; PubMed=11102384;
 RA Kilman R.M., Andolfatto P., Coyne J.A., Depaulis F., Kreitman M.,
 Berry A.J., McCarter J., Wakeley J., Hey J.;
 RT "The Population Genetics of the Origin and Divergence of the
 RT Drosophila simulans Complex Species.";
 RL Genetics 156:1913-1931(2000).
 DR EMBL; AF284453; AAC49486.1; -;
 DR FlyBase; FBgn0016348; Dsinn\Jana.
 FT NON_TER 1
 SQ SEQUENCE 109 AA; 12384 MW; 1118CB3722E56A6F CRC64;

Query Match 43.0%; Score 107; DB 5; Length 109;
 Best Local Similarity 46.5%; Pred. No. 2.4e-07;
 Matches 20; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 YHADIDKVSQDMKOGCCDCECLGGGRISHQSODKKIHVYGS 43
 DB 36 WHADIFEREVEYFKKLGARACPGGRIEHNPKKYLKLVYGS 78

RESULT 13

O8T475 PRELIMINARY; PRT; 119 AA.
 AC O8T475;



GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: May 7, 2003, 19:01:25 ; Search time 11.7064 Seconds
(without alignments)
110.590 Million cell updates/sec

Title: US-09-914-831-5
Sequence: 1 YHADIYDKVSGDMKQKGCDC.....GGRIHQSGDKKHIVGYSM 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Dackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	21.7	149	4	US-09-183-861-86 Sequence 86, Appl
2	54	21.7	149	4	US-09-022-765-86 Sequence 86, Appl
3	53	21.3	445	2	US-08-935-450-4 Sequence 4, Appl
4	51.5	20.7	273	3	US-08-397-411-6 Sequence 6, Appl
5	50.5	20.3	394	1	US-07-621-193A-5 Sequence 5, Appl
6	50.5	20.3	394	1	US-08-018-489C-5 Sequence 5, Appl
7	50	20.1	48	4	US-08-438-745-22 Sequence 22, Appl
8	50	20.1	48	4	PCT-US94-05669A-22 Sequence 22, Appl
9	50	20.1	349	4	US-09-459-774-2 Sequence 2, Appl
10	49.5	19.9	2231	1	US-08-153-799-16 Sequence 16, Appl
11	49.5	19.9	2324	1	US-08-283-857-1 Sequence 1, Appl
12	49.5	19.9	2324	5	PCT-US95-09819-1 Sequence 1, Appl
13	49.5	19.9	2327	6	PCT-US95-09819-1 Patent No. 5455158
14	49.5	19.9	2386	2	US-09-016-366A-12 Sequence 12, Appl
15	49.5	19.9	2446	2	US-08-551-356-2 Sequence 2, Appl
16	49.5	19.9	2446	5	PCT-US93-12687-2 Sequence 2, Appl
17	49	19.7	592	3	US-08-991-813-2 Sequence 2, Appl
18	49	19.7	1211	4	US-09-491-522-5 Sequence 5, Appl
19	48.5	19.5	388	1	US-08-290-448A-80 Sequence 80, Appl
20	48.5	19.5	388	1	US-08-290-448A-80 Sequence 80, Appl
21	48.5	19.5	388	1	US-08-175-069A-80 Sequence 80, Appl
22	48.5	19.5	388	4	US-08-461-939B-80 Sequence 80, Appl
23	48.5	19.5	388	4	US-08-464-000-80 Sequence 80, Appl
24	48	19.3	118	3	US-08-483-304-12 Sequence 12, Appl
25	48	19.3	118	3	US-08-483-304-12 Sequence 12, Appl
26	48	19.3	140	3	US-08-483-304-9 Sequence 9, Appl
27	48	19.3	140	3	US-08-483-304-9 Sequence 9, Appl

28	48	19.3	450	4	US-09-592-891A-14	Sequence 14, Appl
29	48	19.3	846	2	US-07-728-215-33	Sequence 33, Appl
30	48	19.3	846	4	US-08-938-085A-33	Sequence 33, Appl
31	48	19.3	1205	4	US-09-491-522-11	Sequence 11, Appl
32	47.5	19.1	677	3	US-08-480-640A-115	Sequence 115, Appl
33	47.5	19.1	677	3	US-08-480-640A-115	Sequence 115, Appl
34	47.5	19.1	677	3	US-08-295-802-115	Sequence 115, Appl
35	47.5	19.1	677	4	US-08-686-968C-58	Sequence 115, Appl
36	47.5	19.1	677	4	US-08-686-968C-58	Sequence 115, Appl
37	47.5	19.1	677	4	US-08-488-237A-115	Sequence 115, Appl
38	47.5	19.1	677	4	US-08-488-237A-115	Sequence 115, Appl
39	47.5	19.1	677	4	US-08-375-992A-115	Sequence 115, Appl
40	47.5	19.1	677	4	US-08-375-992A-115	Sequence 115, Appl
41	47	18.9	235	4	US-08-438-745-13	Sequence 13, Appl
42	47	18.9	235	4	US-09-219-019-13	Sequence 13, Appl
43	47	18.9	235	5	PCT-US94-05669A-13	Sequence 13, Appl
44	47	18.9	591	3	US-08-965-903B-2	Sequence 2, Appl
45	46.5	18.7	333	4	US-09-230-637-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-09-183-861-86
Sequence 86, Application US/09183861
Patent No. 6365165
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Melo, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, David C.
APPLICANT: Skelly, Yasir A.W.
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,861
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/022,765
FILING DATE: 12-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.42003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-183-861-86
Query Match 21.7%; Score 54; DB 4; Length 149;
Best Local Similarity 35.7%; Pred. No. 4.8;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 3 ADIYKVSQDMQKQGCCEICGGGRISH 30
|::|||::|||::|||
Db 65 AELIDSVLDVCKREAESCDLQGFOLSH 92

RESULT 2

US-09-022-765-86
; Sequence 86, Application US/09022765
; Patent No. 6375955
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,765
; FILING DATE: 12-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.420C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-022-765-86

Query Match 21.7%; Score 54; DB 4; Length 149;
Best Local Similarity 35.7%; Pred. No. 4.8;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 3 ADIYKVSQDMQKQGCCEICGGGRISH 30
|::|||::|||::|||
Db 65 AELIDSVLDVCKREAESCDLQGFOLSH 92

RESULT 3

US-08-935-450-4
; Sequence 4, Application US/08935450
; Patent No. 5977311
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Melija
; APPLICANT: Schulz, Vincent
; TITLE OF INVENTION: 53BP2 COMPLEXES
; FILE REFERENCE: 7934-054
; CURRENT APPLICATION NUMBER: US/08/935,450
; CURRENT FILING DATE: 1997-09-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 445

; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-935-450-4

Query Match 21.3%; Score 53; DB 2; Length 445;
Best Local Similarity 32.1%; Pred. No. 22;
Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 3 ADIYKVSQDMQKQGCCEICGGGRISH 30
|::|||::|||::|||
Db 110 AELVDSVLDVCKREAESCDLQGFOLTH 137

RESULT 4

US-08-397-411-6
; Sequence 6, Application US/08397411
; Patent No. 6129914
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Gingrich, Roger
; APPLICANT: Link, Brian
; APPLICANT: Tso, J. Yun
; TITLE OF INVENTION: Bispecific Antibody Effective to Treat
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,411
; FILING DATE: 01-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/859,583
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 011823-004901
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-397-411-6

Query Match 20.7%; Score 51.5; DB 3; Length 273;
Best Local Similarity 28.6%; Pred. No. 21;
Matches 12; Conservative 6; Mismatches 17; Indels 7; Gaps 1;

QY 2 HADIYKVSQDMQKQGC-----CEICGGGRISHOSQDKK 36
|::|||::|||::|||
Db 203 HKPSNTKVDKVKPKSCDKTHCPCKCPAGSRIRALEYKA 244

RESULT 5

US-07-621-193A-5
; Sequence 5, Application US/07621193A
; Patent No. 5187087

GENERAL INFORMATION:
APPLICANT: Sondermeyer, Paulus Jacobus Antonius
APPLICANT: Claessens, Johannes Antonius Joseph
APPLICANT: Mockett, Albert Philip Adrian
TITLE OF INVENTION: Recombinant herpesvirus of
TITLE OF INVENTION: turkeys and live vector vaccines derived thereof
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Organon Teknika Corporation
ADDRESSEE: Biotechnology Research Institute
STREET: 1330-A Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/621,193A
FILING DATE: 19901130
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Donna Bobrowicz
REGISTRATION NUMBER: 32,196
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301)258-5200
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE: ORF-5
US-07-621-193A-5

Query Match 20.3%; Score 50.5; DB 1; Length 394;
Best Local Similarity 31.7%; Pred. No. 42;
Matches 13; Conservative 6; Mismatches 11; Indels 11; Gaps 1;

QY 11 GDMQKQGCDECELGGR-----ISHQSDKKIHVY 40
DB 115 GDADQKCIYKAVYGGKNGREVYDIKTISHKSIKLIHAY 155

RESULT 6
US-08-018-489C-5
Sequence 5, Application US/08018489C
Patent No. 5470734
GENERAL INFORMATION:
APPLICANT: Sondermeyer, Paulus Jacobus Antonius
APPLICANT: Claessens, Johannes Antonius Joseph
APPLICANT: Mockett, Albert Philip Adrian
TITLE OF INVENTION: Recombinant herpesvirus of
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo Pharma
STREET: 1330-A Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/018,489C
FILING DATE: February 12, 1993
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/621,193
FILING DATE: 30-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Donna Bobrowicz
REGISTRATION NUMBER: 32,196
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301)258-5200
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE: ORF-5
US-08-018-489C-5

Query Match 20.3%; Score 50.5; DB 1; Length 394;
Best Local Similarity 31.7%; Pred. No. 42;
Matches 13; Conservative 6; Mismatches 11; Indels 11; Gaps 1;

QY 11 GDMQKQGCDECELGGR-----ISHQSDKKIHVY 40
DB 115 GDADQKCIYKAVYGGKNGREVYDIKTISHKSIKLIHAY 155

RESULT 7
US-08-438-745-22
Sequence 22, Application US/08438745
Patent No. 6248715
GENERAL INFORMATION:
APPLICANT: Rosenberg, Steven
APPLICANT: Stratton-Thomas, Jennifer
TITLE OF INVENTION: Expression of Urokinase Plasminogen
TITLE OF INVENTION: Activator Inhibitors
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,745
FILING DATE: 10-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/070,153
FILING DATE: 01-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-655-3542
TELEFAX: 510-601-2706
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO

```

; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
US-08-438-745-22

```

Query Match	20.1%	Score 50;	DB 4;	Length 488
Best Local Similarity	36.4%	Pred. No. 4.7;		
Matches	8;	Conservative	5;	Mismatches 9; Indels

```
QY 10 SGDMQKQGCDCCECLGGGRISHQ 31
    | :: : |:| | | |
Db 1 SNELHQVPSNCDCLNGGTCVHQ 22
```

RESULT 8
PCT-US94-05669A-22
; Sequence 22, Application PC/TUS9405669A

APPLICANT: Chiron Corporation
TITLE OF INVENTION: Expression of Urokinase Plasminogen
TITLE OF INVENTION: Activator Inhibitors
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608

Query Match	20.1%;	Score 50;	DB 5;	Length 48;
Best Local Similarity	36.4%;	Pred. No. 4.7;		
Matches	8;	Conservative	5;	Mismatches 9;
				Indels

```
QY 10 SGDMQKQCGDCCECLGGGRISHQ 31
      |::: :|:| | |
Db 1 SNEHQVPSPNCDCCLNGGTCVHQ 22
```

RESULT 9
US-09-459-774-2
; Sequence 2, Application US/09459774
; Patent No. 6297030
; GENERAL INFORMATION:
; APPLICANT: Michael Robert Barnes
; APPLICANT: Tania Tamson Testa
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30193
; CURRENT APPLICATION NUMBER: US/09/459, 774/

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	20.1%;	Score 50;	DB 4;	Length 349;
Best Local Similarity	32.4%;	Pred. No. 43;		
Matches 12;	Conservative 3;	Mismatches 10;	Indels 12;	Gaps 1;

```

Qy      7  DKVSGDM-----QKGGCDECELGGRISHQ 31
          | | | :      | | | | | | : | |
Db      281 DPTGSGVGTGGRACNKTAPOAGSCDLMCCGRYNTHQ 317

```

RESULT 10
US-08-153-799-16
; Sequence 16, Application US/08153799
Data 4 us 576003

APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /note="Human fibronectin"
US-08-153-799-16

Query Match 19.9%; Score 49.5; DB 1; Length 2231;
Best Local Similarity 30.3%; Pred. No. 4e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMOKG-----CDCECLGGGR 27
DB 67 FDKYTGNTYRVGDYTERPKDSMIMDCTCIGAGR 99

RESULT 11
US-08-283-857-1

Sequence 1, Application US/08283857
Patent No. 5792742

GENERAL INFORMATION:

APPLICANT: GOLD, Leslie I.
APPLICANT: ROSTAGNO, Agueda A.

APPLICANT: BARON, Martin
APPLICANT: CAMPBELL, Iain D.

TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NETMARK
STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/283,857
FILING DATE: 01-AUG-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/714,134
FILING DATE: 14-JUN-1991

ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.

REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-283-857-1

Query Match 19.9%; Score 49.5; DB 1; Length 2324;
Best Local Similarity 30.3%; Pred. No. 4.2e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMOKG-----CDCECLGGGR 27
DB 67 FDKYTGNTYRVGDYTERPKDSMIMDCTCIGAGR 99

RESULT 12
PCT-US95-09819-1

Sequence 1, Application PC/TUS9509819
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA

TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK

STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington

STATE: D.C.
COUNTRY: USA

ZIP: 20004
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09819

FILING DATE:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/283,857
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.

REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide

PCT-US95-09819-1

Query Match 19.9%; Score 49.5; DB 5; Length 2324;
Best Local Similarity 30.3%; Pred. No. 4.2e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMOKG-----CDCECLGGGR 27
DB 67 FDKYTGNTYRVGDYTERPKDSMIMDCTCIGAGR 99

RESULT 13
5455158-1

Patent No. 5455158
APPLICANT: VOGEL, TIKVA;LEVANON, AVIGDOR;WEBER, MOSHE M.;

GUY, RACHEL;PANET, AMOS
TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND

USES AND METHODS OF PRODUCING SAME
NUMBER OF SEQUENCES: 20

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/58,241

FILING DATE: 04-MAY-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 526,397
FILING DATE: 21-MAY-1990

APPLICATION NUMBER: 345,952
FILING DATE: 28-APR-1989

APPLICATION NUMBER: 291,951
FILING DATE: 29-DEC-1988

SEQ ID NO: 1;
LENGTH: 2327
5455158-1

Query Match 19.9%; Score 49.5; DB 6; Length 2327;
 Best Local Similarity 30.3%; Pred. No. 4.2e+02;
 Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMQKOG-----CDCECLGGGR 27
 DB 70 FDKYTGNTYRVGDTYERPKDSMIMDCTCIGAGR 102

RESULT 14

US-09-016-366A-12

Sequence 12, Application US/09016366A

Patent No. 5955431

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.

APPLICANT: Huang, Chifu

TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE

TITLE OF INVENTION: INHIBITORS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,366A

FILING DATE: January 30, 1998

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/037,090

FILING DATE: 05-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 36,637

REFERENCE/DOCKET NUMBER: B0801/7093

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

TELEX:

INFORMATION FOR SEQ. ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2386 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-016-366A-12

Query Match 19.9%; Score 49.5; DB 2; Length 2386;
 Best Local Similarity 30.3%; Pred. No. 4.3e+02;
 Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMQKOG-----CDCECLGGGR 27
 DB 98 FDKYTGNTYRVGDTYERPKDSMIMDCTCIGAGR 130

RESULT 15

US-08-551-356-2

Sequence 2, Application US/08551356

Patent No. 5830700

GENERAL INFORMATION:

APPLICANT: Iranil, Meher

TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Zymogenetics, Inc.

STREET: 4225 Roosevelt Way, N.E.

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/551,356

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/349,762

FILING DATE:

APPLICATION NUMBER: US/07/998,271

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Holly, Julie A

REGISTRATION NUMBER: 33-246

REFERENCE/DOCKET NUMBER: 92-26

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-547-8080 ext 322

TELEFAX: 206-548-2329

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2446 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-551-356-2

Query Match 19.9%; Score 49.5; DB 2; Length 2446;
 Best Local Similarity 30.3%; Pred. No. 4.5e+02;
 Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMQKOG-----CDCECLGGGR 27
 DB 98 FDKYTGNTYRVGDTYERPKDSMIMDCTCIGAGR 130

Search completed: May 7, 2003, 19:08:39
 Job time: 13.7064 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2003, 19:04:05 ; Search time 10.6972 Seconds

(without alignments)
378.521 Million cell updates/sec

Title: US-09-914-831-5

Perfect score: 249
Sequence: 1 YHADIYDKVSGDMQKQCDC.....GGRISHOSDKKIHYGYSM 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55.5	22.3	430	US-09-741-669-442	Sequence 442, App
2	54	21.7	149	US-09-991-496-86	Sequence 86, Appl
3	54	21.7	149	US-09-874-923-86	Sequence 86, Appl
4	53	21.3	351	US-09-864-761-33332	Sequence 33332, A
5	52	20.9	374	US-09-815-242-10571	Sequence 10571, A
6	52	20.9	494	US-09-764-864-1235	Sequence 1235, App
7	51	20.5	280	US-10-101-464A-511	Sequence 511, App
8	51	20.5	998	US-10-101-464A-514	Sequence 914, App
9	50	20.1	98	US-09-981-876-260	Sequence 260, App
10	50	20.1	98	US-09-148-545-260	Sequence 260, App
11	50	20.1	349	US-09-903-817-2	Sequence 260, App
12	49.5	19.9	164	US-09-925-301-1270	Sequence 1270, App
13	49.5	19.9	231	US-09-925-302-548	Sequence 548, App
14	49.5	19.9	259	US-09-940-235-4	Sequence 548, App
15	49.5	19.9	2386	US-09-961-403-1	Sequence 4, Appl
16	49	19.7	109	US-09-764-891-3528	Sequence 3528, App
17	49	19.7	140	US-09-739-254-76	Sequence 76, Appl
18	49	19.7	140	US-09-904-615-76	Sequence 76, Appl
19	49	19.7	195	US-09-925-302-717	Sequence 717, App

20	49	19.7	245	10	US-09-796-089-9	Sequence 9, Appl
21	49	19.7	246	10	US-09-739-254-143	Sequence 143, App
22	49	19.7	246	10	US-09-904-615-143	Sequence 143, App
23	49	19.7	3479	9	US-10-123-155-123	Sequence 123, App
24	48.5	19.5	397	9	US-09-847-208-17	Sequence 17, Appl
25	48	19.3	450	9	US-09-969-844-14	Sequence 14, Appl
26	48	19.3	708	9	US-10-184-644-211	Sequence 211, App
27	48	19.3	708	9	US-10-184-634-211	Sequence 211, App
28	48	19.3	823	9	US-10-123-155-379	Sequence 379, App
29	48	19.3	846	9	US-10-072-841-33	Sequence 33, Appl
30	48	19.3	969	10	US-09-321-987B-5	Sequence 5, Appl
31	48	19.3	1076	9	US-10-184-644-191	Sequence 191, App
32	48	19.3	1076	9	US-10-184-634-191	Sequence 191, App
33	48	19.3	1103	9	US-10-184-644-189	Sequence 189, App
34	48	19.3	1103	9	US-10-184-634-189	Sequence 189, App
35	48	19.3	1321	9	US-10-184-644-371	Sequence 371, App
36	48	19.3	1321	9	US-10-184-634-371	Sequence 371, App
37	48	19.3	1471	9	US-10-123-155-281	Sequence 281, App
38	48	19.3	1503	9	US-10-123-155-363	Sequence 363, App
39	48	19.3	1869	9	US-10-123-155-269	Sequence 269, App
40	48	19.3	1942	9	US-10-123-155-515	Sequence 515, App
41	48	19.3	1985	9	US-10-123-155-143	Sequence 143, App
42	48	19.3	2213	9	US-10-184-644-549	Sequence 549, App
43	48	19.3	2213	9	US-10-184-634-549	Sequence 549, App
44	48	19.3	2380	9	US-10-184-644-597	Sequence 597, App
45	48	19.3	2380	9	US-10-184-634-597	Sequence 597, App

ALIGNMENTS

```

RESULT 1
US-09-741-669-442
; Sequence 442, Application us/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELIURA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 442
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-442

Query Match 22.3% Score 55.5; DB 10; Length 430;
Best Local Similarity 35.0% Pred. NO. 14;
Matches 14; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

OY 1 YHADIYDKVSGDMQKQCDCGGRISHOSDKKIHY 40
DB 53 FKMDHEKVIILQKMKRIYSLPSGKITKFT-DRKIKHY 91

RESULT 2
US-09-991-496-86
; Sequence 86, Application us/09991496
; Patent No. US20020169285A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.

```

APPLICANT: Bhatia, Ajay
APPLICANT: Cole, Rhea
APPLICANT: Probst, Peter
APPLICANT: Brannon, Mark
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
FILE REFERENCE: 210121.420C9
CURRENT APPLICATION NUMBER: US/09/991,496
NUMBER OF SEQ ID NOS: 137
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 86
LENGTH: 149
TYPE: PRT
ORGANISM: Leishmania major
FEATURE:
NAME/KEY: VARIANT
LOCATION: 8, 109, 112, 136
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-991-496-86

Query Match 21.7%; Score 54; DB 9; Length 149;
Best Local Similarity 35.7%; Pred. No. 7;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

OY 3 ADIDYVSGDMQKGGCCGCGGGRISH 30
DB 65 AELIDSVLDVCKREAESCDCLQGFOLSH 92

US-09-874-923-86
Sequence 86, Application US/09874923
Patent No. US20020081320A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, David C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Bhatia, Ajay
APPLICANT: Cole, Rhea
APPLICANT: Probst, Peter
APPLICANT: Brannon, Mark
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
FILE REFERENCE: 210121.420C8
CURRENT APPLICATION NUMBER: US/09/874,923
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 86
LENGTH: 149
TYPE: PRT
ORGANISM: Leishmania major
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(149)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-874-923-86

Query Match 21.7%; Score 54; DB 10; Length 149;
Best Local Similarity 35.7%; Pred. No. 7;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

OY 3 ADIDYVSGDMQKGGCCGCGGGRISH 30
DB 65 AELIDSVLDVCKREAESCDCLQGFOLSH 92

RESULT 4
US-09-864-761-33332
Sequence 33332, Application US/09864761

Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chan, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acomica-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO 33332
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AB023051.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 51
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 51
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1e+02
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 30
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 23
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 15
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 17
OTHER INFORMATION: EST_HUMAN HIT: BE745540.1, EVALUATE 0.00e+00
OTHER INFORMATION: EST_HUMAN HIT: A1815647.1, EVALUATE 1.00e-125
OTHER INFORMATION: SWISSPROT HIT: P05218, EVALUATE 0.00e+00
US-09-864-761-33332

Query Match 21.3%; Score 53; DB 10; Length 351;
Best Local Similarity 32.1%; Pred. No. 25;
Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

OY 3 ADIYKVSQDMQKQGCCECLGGGRISH 30
 DB 17 AELVDSVLDVVRKKEAESDCJGQFOLTH 44

RESULT 5

US-09-815-242-10571
 ; Sequence 10571, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10571
 ; LENGTH: 374
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-815-242-10571

Query Match 20.9%; Score 52; DB 10; Length 374;
 Best Local Similarity 62.5%; Pred. No. 36;
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 21 ECLGGGRISHQSDKK 36
 DB 354 ECLGGGLIDRAYODEK 369

RESULT 6

US-09-764-864-1235
 ; Sequence 1235, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT223
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; PRIOR FILING DATE: 2001-01-17
 ; PRIOR APPLICATION data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1235
 ; LENGTH: 494
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:

NAME/KEY: SITE
 LOCATION: (14)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (488)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (493)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (494)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-764-864-1235

Query Match 20.9%; Score 52; DB 10; Length 494;
 Best Local Similarity 39.1%; Pred. No. 49;
 Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 22 CLGGGRISHQSDKKTHVGYSM 44
 DB 407 CTNMAEVRHEIDQCTHLVYISL 429

RESULT 7

US-10-101-464A-511
 ; Sequence 511, Application US/10101464A
 ; Publication No. US20030046728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Strabala, Timothy
 ; APPLICANT: Nieuwenhuizen, Nicolaas
 ; APPLICANT: Higgins, Colleen M.
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells
 ; FILE OF INVENTION: and their use in the Modification of Plant Cell Signaling
 ; FILE REFERENCE: 11000.1020C2
 ; CURRENT APPLICATION NUMBER: US/10/101,464A
 ; PRIOR FILING DATE: 2002-03-18
 ; PRIOR APPLICATION NUMBER: 09/704,302
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 09/228,986
 ; PRIOR FILING DATE: 1995-01-12
 ; PRIOR APPLICATION NUMBER: 60/162,866
 ; PRIOR FILING DATE: 1999-11-01
 ; PRIOR APPLICATION NUMBER: PCT/US00/00724
 ; PRIOR FILING DATE: 2000-01-11
 ; NUMBER OF SEQ ID NOS: 989
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 511
 ; LENGTH: 280
 ; TYPE: PRT
 ; ORGANISM: Eucaalyptus grandis
 ; US-10-101-464A-511

Query Match 20.5%; Score 51; DB 9; Length 280;
 Best Local Similarity 28.0%; Pred. No. 36;
 Matches 14; Conservative 9; Mismatches 17; Indels 10; Gaps 1;

OY 3 ADIYKVSQDMQKQGCCECLGGGRISHQ-----SODKKTHVGY 42
 DB 95 SESHVEKRGHYDDGFDAEVRTLGRIRKNIVKIMCSTTRDCKLVY 144

RESULT 8

US-10-101-464A-914
 ; Sequence 914, Application US/10101464A
 ; Publication No. US20030046728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Strabala, Timothy
 ; APPLICANT: Nieuwenhuizen, Nicolaas
 ; APPLICANT: Higgins, Colleen M.
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells
 ; FILE OF INVENTION: and their use in the Modification of Plant Cell Signaling
 ; CURRENT APPLICATION NUMBER: US/10/101,464A

;; CURRENT FILING DATE: 2002-03-18
;; PRIOR APPLICATION NUMBER: 09/704,302
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 09/228,986
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/162,866
;; PRIOR FILING DATE: 1999-11-01
;; PRIOR APPLICATION NUMBER: PCI/US00/00724
;; PRIOR FILING DATE: 2000-01-11
;; NUMBER OF SEQ ID NOS: 989
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 914
;; LENGTH: 998
;; TYPE: PRT
;; ORGANISM: Eucalyptus grandis
US-10-101-464A-914

Query Match 20.5%; Score 51; DB 9; Length 998;
Best Local Similarity 28.0%; Pred. No. 1.5e+02;
Matches 14; Conservative 9; Mismatches 17; Indels 10; Gaps 1;

QY 3 ADIYRVSDMKGCCDCECLGGGRISHQ-----SQDKIHVGY 42
Db 723 SESHVDEKSPVDDGDAEYKTLGRHKNVYKLMCSCTTRDCKLVEY 772

RESULT 9
US-09-981-876-260
; Sequence 260, Application US/09981876
; Patent No. US2002016469A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592

;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,581
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,584
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,500
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,587
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,492
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,598
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,613
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,582
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,596
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,612
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,632
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/047,601
;; PRIOR FILING DATE: 1997-05-23
;; PRIOR APPLICATION NUMBER: 60/043,580
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,568
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,314
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,569
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,311
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,671
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,674
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,669
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,312
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,313
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,672
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/043,315
;; PRIOR FILING DATE: 1997-04-11
;; PRIOR APPLICATION NUMBER: 60/048,974
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/056,886
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,877
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,889
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,893
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,630
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,878
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,662
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,872
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,882
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,637
;; PRIOR FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: 60/056,903
;; PRIOR FILING DATE: 1997-08-22

PRIOR APPLICATION NUMBER: 60/056,888
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,879
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,880
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,894
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,911
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,636
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,874
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,910
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,864
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,631
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,845
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,892
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/047,595
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/057,761
PRIOR FILING DATE: 05-Sep-1997
PRIOR APPLICATION NUMBER: 60/047,599
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,588
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,585
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,586
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,590
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,594
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,589
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,593
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,614
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,578
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,576
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/047,501
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,670
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/056,632
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,664
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,876
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,881
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,909
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,875
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,862
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,887
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,908
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/048,964

PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/057,650
PRIOR FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: 60/056,884
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 260
LENGTH: 98

Query Match 20.1%; Score 50; DB 9; Length 98;
Best Local Similarity 32.4%; Pred. No. 16;
Matches 12; Conservative 3; Mismatches 10; Indels 12; Gaps 1;

7 DKVSGDM-----OKGCDCECGGGGRISHQ 31
Db 36 DPTVSGVGTGRACNKTAPQASGCDLMCCGRGYTHQ 72

RESULT 10
US-09-148-545-260
Sequence 260, Application US/09148545
Publication No. US20030327132A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: P2001P1
CURRENT APPLICATION NUMBER: US/09/148,545
CURRENT FILING DATE: 1998-09-04
EARLIER APPLICATION NUMBER: PCT/US98/04482
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,161
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23

1	EARLIER FILING DATE: 1997-08-22
2	EARLIER APPLICATION NUMBER: 60/056, 911
3	EARLIER FILING DATE: 1997-08-22
4	EARLIER APPLICATION NUMBER: 60/056, 636
5	EARLIER FILING DATE: 1997-08-22
6	EARLIER APPLICATION NUMBER: 60/056, 874
7	EARLIER FILING DATE: 1997-08-22
8	EARLIER APPLICATION NUMBER: 60/056, 910
9	EARLIER FILING DATE: 1997-08-22
10	EARLIER APPLICATION NUMBER: 60/056, 864
11	EARLIER FILING DATE: 1997-08-22
12	EARLIER APPLICATION NUMBER: 60/056, 631
13	EARLIER FILING DATE: 1997-08-22
14	EARLIER APPLICATION NUMBER: 60/056, 845
15	EARLIER FILING DATE: 1997-08-22
16	EARLIER APPLICATION NUMBER: 60/056, 892
17	EARLIER FILING DATE: 1997-08-22
18	EARLIER APPLICATION NUMBER: 60/047, 595
19	EARLIER FILING DATE: 1997-05-23
20	EARLIER APPLICATION NUMBER: 60/047, 594
21	EARLIER FILING DATE: 1997-05-23
22	EARLIER APPLICATION NUMBER: 60/047, 589
23	EARLIER FILING DATE: 1997-05-23
24	EARLIER APPLICATION NUMBER: 60/047, 585
25	EARLIER FILING DATE: 1997-05-23
26	EARLIER APPLICATION NUMBER: 60/047, 586
27	EARLIER FILING DATE: 1997-05-23
28	EARLIER APPLICATION NUMBER: 60/047, 580
29	EARLIER FILING DATE: 1997-05-23
30	EARLIER APPLICATION NUMBER: 60/047, 594
31	EARLIER FILING DATE: 1997-05-23
32	EARLIER APPLICATION NUMBER: 60/047, 589
33	EARLIER FILING DATE: 1997-05-23
34	EARLIER APPLICATION NUMBER: 60/047, 593
35	EARLIER FILING DATE: 1997-05-23
36	EARLIER APPLICATION NUMBER: 60/047, 614
37	EARLIER FILING DATE: 1997-05-23
38	EARLIER APPLICATION NUMBER: 60/043, 578
39	EARLIER FILING DATE: 1997-04-11
40	EARLIER APPLICATION NUMBER: 60/043, 576
41	EARLIER FILING DATE: 1997-04-11
42	EARLIER APPLICATION NUMBER: 60/047, 501
43	EARLIER FILING DATE: 1997-05-23
44	EARLIER APPLICATION NUMBER: 60/043, 670
45	EARLIER FILING DATE: 1997-04-11
46	EARLIER APPLICATION NUMBER: 60/056, 633
47	EARLIER FILING DATE: 1997-08-22
48	EARLIER APPLICATION NUMBER: 60/056, 664
49	EARLIER FILING DATE: 1997-08-22
50	EARLIER APPLICATION NUMBER: 60/056, 876
51	EARLIER FILING DATE: 1997-08-22
52	EARLIER APPLICATION NUMBER: 60/056, 881
53	EARLIER FILING DATE: 1997-08-22
54	EARLIER APPLICATION NUMBER: 60/056, 909
55	EARLIER FILING DATE: 1997-08-22
56	EARLIER APPLICATION NUMBER: 60/056, 875
57	EARLIER FILING DATE: 1997-08-22
58	EARLIER APPLICATION NUMBER: 60/056, 863
59	EARLIER FILING DATE: 1997-08-22
60	EARLIER APPLICATION NUMBER: 60/056, 884
61	EARLIER FILING DATE: 1997-08-22
62	EARLIER APPLICATION NUMBER: 60/056, 908
63	EARLIER FILING DATE: 1997-08-22
64	EARLIER APPLICATION NUMBER: 60/048, 964
65	EARLIER FILING DATE: 1997-06-06
66	EARLIER APPLICATION NUMBER: 60/057, 656
67	EARLIER FILING DATE: 1997-09-05
68	EARLIER APPLICATION NUMBER: 60/056, 884
69	EARLIER FILING DATE: 1997-08-22
70	NUMBER OF SEQ ID NOS: 280
71	SOFTWARE: Patent In. Ver. 2.0

SEQ ID NO 260
LENGTH: 98

Query Match
Best Local Similarity 20.1%; Score 50; DB 9; Length 98;
Matches 12; Conservative 3; Mismatches 10; Indels 12; Gaps 1;

QY 7 DKVSGDM-----QKQCCDECTGGGRISHQ 31
DB 36 DPTGSGVTGGRACNKTAPQASGCDLMCCGRGNTHTQ 72

RESULT 11

US-09-903-817-2
Sequence 2, Application US/09903817
Patent No. US20010055790A1
GENERAL INFORMATION:
APPLICANT: BARRES, Michael Robert
APPLICANT: TESTA, Tania Tamson
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30193-D1
CURRENT APPLICATION NUMBER: US/09/903,817
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: UK 9828419.3
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 09/459,774
PRIOR FILING DATE: 1999-12-13
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 349
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-903-817-2

Query Match
Best Local Similarity 32.4%; Score 50; DB 10; Length 349;
Matches 12; Conservative 3; Mismatches 10; Indels 12; Gaps 1;

QY 7 DKVSGDM-----QKQCCDECTGGGRISHQ 31
DB 281 DPTGSGVTGGRACNKTAPQASGCDLMCCGRGNTHTQ 317

RESULT 12

US-09-925-301-1270
Sequence 1270, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: ROSEN, et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1270
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (13)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (138)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (152)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (161)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (164)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-301-1270

Query Match
Best Local Similarity 19.9%; Score 49.5; DB 10; Length 164;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMQKQ-----CDCECTGGGR 27
DB 17 FDKYTGNTYRGDTPRPRKDSMIMDCTCIGAGR 49

RESULT 13
US-09-925-302-548
Sequence 548, Application US/09925302
Patent No. US2002004941A1
GENERAL INFORMATION:
APPLICANT: ROSEN, et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 548
LENGTH: 231
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (44)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (205)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (212)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (226)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-548

Query Match
Best Local Similarity 19.9%; Score 49.5; DB 10; Length 231;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 6 YDKVSGDMQKQ-----CDCECTGGGR 27
DB 127 FDKYTGNTYRGDTPRPRKDSMIMDCTCIGAGR 159

RESULT 14
US-09-940-235-4
Sequence 4, Application US/09940235
Publication No. US20030059921A1
GENERAL INFORMATION:

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (138)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (152)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (161)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (164)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1270

```

APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammara
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 259
TYPE: PRT
ORGANISM: Homo sapiens
US-09-940-235-4

```

```

Query Match      19.9%; Score 49.5; DB 9; Length 259;
Best Local Similarity 30.3%; Pred. No. 53;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

```

```

OY      6 YDKVSGDMOKG-----CDCECLGGGR 27
      11111111111111111111111111111111
DB      67 FDKYTGNTYRVGDTYRPRKDSMTMDCTCIGAGR 99

```

```

RESULT 15
US-09-961-403-1
; Sequence 1, Application US/09961403
; Publication No. US2003007589A1
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPF, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

```

```

Query Match      19.9%; Score 49.5; DB 9; Length 2386;
Best Local Similarity 30.3%; Pred. No. 6.1e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

```

```

OY      6 YDKVSGDMOKG-----CDCECLGGGR 27
      11111111111111111111111111111111
DB      98 FDKYTGNTYRVGDTYRPRKDSMTMDCTCIGAGR 130

```

Search completed: May 7, 2003, 19:18:51
 Job time : 12.6972 secs